

Struktur und Dynamik kritischer boolescher Zufallsnetzwerke als Modelle der genetischen Regulation

Vom Fachbereich Physik
der Technischen Universität Darmstadt

zur Erlangung des Grades
eines Doktors der Naturwissenschaften
(Dr. rer. nat.)

genehmigte

Dissertation

angefertigt von

Dipl.-Phys. Viktor Kaufman
aus Lviv

Referent: Prof. Dr. B. Drossel
Korreferent: Prof. Dr. M. Porto
Tag der Einreichung: 18.10.2006
Tag der Prüfung: 11.12.2006

Darmstadt 2006
D17

Die Biene fliegt auf eine gute Blume.

Abstract

Genetic regulation is responsible for much of the complexity in living matter. In this thesis, I study one of the oldest and most established generic dynamical models of genetic regulation. In the deterministic *random Boolean network model* (RBN), each node is assigned a Boolean function of its two incoming links, the nodes represent either expressed or not expressed genes, the links represent genetic regulations, the states of the network are the expression patterns. For a special choice of functions, the model is *at its critical point*. The networks are then both highly susceptible to local changes and structured, since some nodes are not affected by the changes. This biologically interesting behaviour is known as the edge of chaos.

The *main objective of this work* was to *understand the dynamics* of critical RBN in the limit of large network sizes *and* to identify the dynamically developing *structure* of the model, together with its *implications* for the dynamical attractors.

The thesis is divided into two parts. After an overview of Boolean models and their applications to genetic regulatory systems, the *structure* of critical RBN is being *studied in the first part* of the work. This study leads eventually to an intuitive understanding of many dynamical properties that could not be previously accessed by numerical simulations. The most important analytical structural findings concern the so called relevant nodes, which alone determine the dynamical fate of the model. They underlie a distinct universal scaling behaviour, which is being characterised using methods of statistical mechanics. Relevant nodes are organised in independent relevant components. One can study in detail simpler relevant components that both act as generic models of real genetic regulatory circuits, and form building blocks of more complex networks. I analyse, for example, their distribution in the ensemble of networks.

In the second part of the work I turn to *dynamical properties* of the model. Simpler components show characteristic behaviour patterns. Already for simpler components, the average number and length of attractors, somewhat unexpectedly, increases exponentially with system size. The same statement holds for general critical RBN. Numerical simulations of these and other properties agree well with corresponding analytical results, albeit the simulations face serious limitations.

An important insight for this work was that one has to understand interconnections between structure and dynamics in complex systems, and that a well - balanced combination of analytical and numerical methods is crucial for success. This insight leads at least for RBN to an intuitive understanding of the model, which is in the meantime widely used in many disciplines.

Zusammenfassung

Das XXI Jahrhundert wird wahrscheinlich im Zeichen der Erforschung der lebenden Materie stehen. Von rapide steigenden Aktivitäten auf diesem Gebiet zeugen die 260 vollständig sequenzierten und 1100 sich in Arbeit befindenden Genome (Stand März 2005).

Von der Struktur der Genome zur Funktion der Organismen liegt ein langer Weg. Immense Datenmengen suchen ihre Erforscher, während die Komplexität der Organismen nicht annähernd erfasst ist. Ein Lichtblick in diesem Wirrwarr verspricht die Einsicht, dass universelle, von den biochemischen und anderen Details unabhängige Mechanismen diese Komplexität erzeugen. Das Aufstellen und die Analyse der entsprechenden Modelle haben unter anderem das Interesse der Physiker geweckt.

In dieser Arbeit beschäftige ich mich mit einem der ältesten und zugleich erfolgreichsten generischen dynamischen Modelle der für das komplexe Verhalten von Organismen maßgeblichen genetischen Regulation. Gemeint ist das deterministische *Modell der booleschen Zufallsnetzwerke* (RBN) mit zwei eingehenden Verbindungen und einer booleschen Funktion pro Knoten. Die Knoten repräsentieren die Gene, die entweder aktiv oder inaktiv sein können, die Verbindungen repräsentieren die Genregulationen, der Zustand des Netzwerks beschreibt das momentane Aktivitätsmuster.

Das Modell befindet sich bei einer bestimmten Wahl der Funktionen *am kritischen Punkt*. An diesem Punkt wirkt sich eine lokale Zustandsänderung im Mittel auf eine nicht trivial mit der Netzwerkgröße skalierende Anzahl der Knoten aus. Folglich reagiert das Netzwerk empfindlich auf Änderungen, ist aber schon deswegen geordnet und nicht chaotisch, weil Änderungen nicht alle Knoten betreffen. Man spricht von der Ordnung an der Grenze zum Chaos. Um dieses biologisch interessante Verhalten zu zeigen, müssen die Knoten, gegebenenfalls erst im Laufe der Dynamik, nach ihrer Beteiligung an der Ausbildung der Attraktoren der Dynamik unterschieden werden können. Eine gewisse Struktur muss für das Modell charakteristisch sein.

Es war das *Hauptziel dieser Arbeit, die Dynamik* der kritischen RBN mit zwei Eingängen pro Knoten im Grenzfall großer Netzwerke zu *verstehen, insbesondere die Struktur* zu untersuchen *und damit verbundene Folgerungen für die Dynamik* zu identifizieren.

Nach einem Überblick über das Thema der booleschen Modelle und der Modellierung der genetischen Regulation wird *im ersten Teil dieser Arbeit die Struktur* der kritischen RBN untersucht. Die Beschreibung der Struktur hat sich als der Heilige

Gral für ein intuitives Verständnis des Modells erwiesen. Viele dynamische Eigenschaften, die in der Literatur seit Langem mit mäßigem Erfolg numerisch untersucht wurden, lassen sich direkt aus den Strukturkenntnissen folgern. Zu den wichtigsten strukturellen Erkenntnissen gehören die analytischen Ergebnisse für so genannte relevante Knoten. Unter allen Knoten bestimmen die relevanten Knoten die Dynamik. Für sie werden unter anderem die Anzahl und die Wahrscheinlichkeitsverteilung bestimmt. Man findet Universalität im Skalenverhalten mit der Netzwerkgröße. Dabei kommen Methoden der statistischen Physik und der mathematischen Statistik zum Einsatz.

Relevante Knoten sind in unabhängige relevante Komponenten organisiert. Die einfacheren relevanten Komponenten können losgelöst vom Netzwerkensemble betrachtet werden. Sie sind interessant, weil sie als generische Modelle der echten genetischen regulatorischen Teilnetzwerke oder als Bausteine der komplexeren Netzwerke angesehen werden können. Ihre relativen Gewichte und Verteilungen im Netzwerkensemble der kritischen RBN werden bestimmt.

Biologisch relevanten kanalisierenden Netzwerken kommt viel Aufmerksamkeit zu. Es stellt sich heraus, dass sie in fast jeder Hinsicht einen Spezialfall der kritischen RBN bilden.

Im zweiten Teil der Arbeit wird *die Dynamik* der kritischen RBN unter die Lupe genommen. Zuerst werden charakteristische Verhaltensmuster in den einfacheren Netzwerkkomponenten gefunden. Es zeigt sich anschließend, dass bereits einfache Komponenten sehr viele sehr lange Attraktoren haben und damit ein sehr kompliziertes Verhalten zeigen.

Es verwundert dann nicht, dass man mit Hilfe der Erkenntnisse aus dem ersten Teil der Arbeit für die allgemeinen kritischen Netzwerke folgert, dass ihre mittlere Anzahl und Länge der Attraktoren mit der Netzwerkgröße exponentiell ansteigt. Numerische Simulationen werden daraufhin untersucht, diese und andere Eigenschaften zuverlässig wiedergeben zu können.

Eine möglicherweise triviale Einsicht aus dieser Arbeit besteht darin, dass es sinnvoll ist, komplexe Systeme im Bezug auf das Zusammenspiel der Struktur und Dynamik mit Hilfe einer passenden Kombination der analytischen und numerischen Methoden zu studieren. Auf jeden Fall führt diese Strategie zum intuitiven Verständnis des RBN - Modells aus dem Jahr 1969. Dieses Modell hat seinen Platz in vielen Wissenschaftszweigen gefunden. Im biologischen Kontext kann es generisch Vorgänge wie Zelldifferenzierung, Multistabilität, Oszillation, Homöostase, Signaltransduktion und andere beschreiben.

Inhaltsverzeichnis

1. Einleitung	1
1.1. Beispiele komplexer Systeme	6
1.2. Komplexität	8
1.3. Selbstorganisierte Kritikalität	10
2. Überblick: Boolesche Zufallsnetzwerke (RBN)	13
2.1. Einführung	13
2.2. Anwendungsgebiet	14
2.2.1. Überblick	14
2.2.2. Modellierung der genetischen Regulation	15
Nicht boolesche Modelle	16
Hybride Modelle	17
2.2.3. Ausgewählte Anwendungen	19
2.2.4. Robustheit	20
2.2.5. Module, Motive und Komponenten	21
2.3. Modelltypen	23
2.3.1. Netzwerktopologie	23
2.3.2. Aktualisierung	25
Knotenfunktionen	25
Verteilung der Knotenfunktionen, Symmetrie	28
Abfolge der Aktualisierungen	29
2.4. Struktur und Dynamik	32
I. Struktur der RBN mit $K = 2$	35
3. Einfache Netzwerkkomponenten	39
3.1. Einfache Schleifen	39
3.2. Komponenten mit Verzweigungen	40
3.2.1. Kleine Anzahl von Verzweigungen	41
3.2.2. Anzahl der komplexen Komponenten	42

4. Kritische RBN mit $K = 2$	45
4.1. Nicht gefrorene Knoten	47
4.1.1. Stochastischer Prozess zur Bestimmung des gefrorenen Kerns	48
4.1.2. Mean - Field - Betrachtung	49
4.1.3. Skalierung unter Berücksichtigung der Fluktuationen	51
Fokker - Planck - Gleichung	51
Skalenverhalten	53
Wahrscheinlichkeitsverteilung	54
4.1.4. Nicht gefrorene Knoten mit zwei relevanten Eingängen . . .	57
4.2. Nicht gefrorene Knoten in kanalisierenden Netzwerken	57
4.2.1. Stabilisierende Schleifen	59
4.2.2. Skalierung der Anzahl gefrorener Knoten	59
4.3. Relevante Knoten	62
4.3.1. Stochastischer Prozess zur Bestimmung relevanter Knoten .	63
4.3.2. Skalierung der Anzahl relevanter Knoten	67
4.3.3. Relevante Knoten mit zwei relevanten Eingängen	69
4.4. Analogie zwischen $K = 2$ und $K = 1$ Netzwerken	70
4.4.1. Größe stabilisierender Schleifen	72
4.5. Relevante Komponenten	73
4.5.1. Verteilung der komplexen relevanten Komponenten	74
4.5.2. Anzahl der relevanten Komponenten	77
 II. Dynamik der RBN mit $K = 2$	 83
5. Einfache Netzwerkkomponenten	87
5.1. Einfache Schleifen	87
5.2. Komponenten mit Verzweigungen	88
5.2.1. Eine Verzweigung: Zwei Schleifen	89
Fall 1: Primzahlen N_1 und N_2	90
Fall 2: Gleich große $N_1 = N_2 \equiv N$	93
Fall 3: Allgemeine N_1 und N_2	94
5.2.2. Eine Verzweigung: Querverbindung	98
Fall 1: $f_\Sigma = f_{14}$	99
Fall 2: $f_\Sigma = f_4$	99
Fall 3: $f_\Sigma = f_1$	100
Fall 4: $f_\Sigma = f_{11}$	102
Fall 5: $f_\Sigma = f_9$	104
Ergänzende Ergebnisse	106
5.2.3. Anwendbarkeit der Simulationen	106

6. Kritische RBN mit $K = 2$	109
6.1. Nicht auf allen Attraktoren gefrorene Knoten	110
6.2. Anzahl und Länge der Attraktoren	113
6.2.1. Mittlere Anzahl der Attraktoren der Länge L	115
6.2.2. Verteilung der Attraktorlängen	116
6.2.3. Mittlere Anzahl der Attraktoren	119
7. Ausblick	123
A. Entwicklung der Theorie der Komplexität	127
B. Anzahl der Zyklen für zwei Schleifen mit Zwischenverbindung	141
B.1. Ergänzende Beispiele	141
B.2. Begründung der Formeln (5.8) und (5.9)	141
C. Ergänzende Ergebnisse für Schleifen mit einer Querverbindung	145
Literaturverzeichnis	147

1. Einleitung

Kann man die Natur mit den Gesetzen der Physik begreifen, so dass auch universell gültige Vorhersagen für komplexe Vorgänge gemacht werden können? Wenn man sich diese Frage stellt, wird schnell klar, dass die Lücken im Verständnis sehr groß sind. Sollte man dann als Physiker das Feld nicht einfach den anderen Wissenschaften mit ihren eigenen Abstraktionsebenen überlassen, oder lohnt es sich nach physikalisch begründeten Gesetzmäßigkeiten oder sogar Gesetzen zu suchen, die einige Lücken schließen? Die Quintessenz der neuen Erkenntnisse der letzten Jahrzehnte liegt darin, dass breit anwendbare, auf leicht realisierbaren bzw. empirisch plausiblen Annahmen basierende, kontrollierbare Modelle von komplexen Vorgängen ein reichhaltiges und gleichzeitig ein nach bestimmten statistischen Regeln geordnetes Verhalten zeigen, was die außerordentliche Vielfalt in der Natur verständlicher erscheinen lässt.

Insbesondere Leben baut auf emergente Vielfalt und Ordnung fern vom Gleichgewicht auf. Emergente Vielfalt und Ordnung findet man auch in der unbelebten Natur überall. Das Studium dieser allgegenwärtigen Phänomene wirft Fragen im Rahmen der Ontologie¹ und der Erkenntnistheorie² auf. Intensive Untersuchungen dieser komplexen interdisziplinären Fragestellungen führten in den letzten etwa 40 Jahren zu einem vermutlich revolutionären Fortschritt für unser Weltverständnis. Gleichzeitig deutet sich ein Paradigmenwechsel beim anwendungsorientierten Studium von realen komplexen Systemen fern vom Gleichgewicht an. Der Trend geht zur ganzheitlichen statt der mechanistischen Problembetrachtung, zum Studium der Organisation bzw. Ordnung im System aus verschiedenen komplementären Blickwinkeln.

Die booleschen Zufallsnetzwerke, ein Modell, das an dieser Entwicklung von Anfang an Teil hat, ist eines der meist zitierten in den letzten Jahrzehnten. Neue Ergebnisse für dieses Modell³ werden in dieser Arbeit präsentiert.

Kurz gefasst, verstehe ich unter dem *Standardmodell* Netzwerke aus N Knoten mit zwei zufällig gewählten eingehenden Verbindungen pro Knoten. Wie in der Graphentheorie verstehe ich unter einem *Netzwerk* einen gerichteten Graphen ohne Mehrfachkanten. Der *Graph* besteht per Definition aus den Knoten und Kanten zwischen den Knoten. Jede Kante im Netzwerk verbindet einen Ausgang von einem Knoten mit einem Eingang von einem Knoten, auch Selbstverbindungen werden zugelassen.

¹Ontologie fragt nach den Grundstrukturen alles Seienden und dem Sein.

²Erkenntnislehre fragt nach der Möglichkeit des Erkenntnisgewinns und nach Erkenntnisweisen.

³Das Modell wurde ursprünglich von Stuart Kauffman eingeführt mit dem Ziel genetische regulatorische Netzwerke zu beschreiben.

1. Einleitung

Jedem Knoten wird ein boolescher Anfangswert zugewiesen, 0 oder 1, und eine boolesche Funktion, die *Knotenfunktion*, die auf eine deterministische Weise festlegt, welchen booleschen Wert der Knoten in Abhängigkeit von den Werten der *Eingänge* (der Knoten, die durch eine gerichtete Kante zu dem gewählten verbunden sind) annimmt, wenn eine *Aktualisierung* von diesem Knoten durchgeführt wird. Die Angabe aller Knotenwerte bezeichnet man als *Netzwerkzustand*. Die Angabe der Verbindungen im Netzwerk (der *Netzwerktopologie*) und der Funktionen der Knoten legt eine *Netzwerkrealisierung* fest. Alle Netzwerkrealisierungen bilden den *Netzwerk - Konfigurationsraum*, im Folgenden auch *Netzwerk - Ensemble* genannt, das vom Phasenraum der Netzwerkzustände zu unterscheiden ist.

Eine Netzwerkrealisierung wird zu einem diskreten dynamischen System durch die Angabe der *Anfangswerte* und die anschließende wiederholte Aktualisierung von allen Knoten. Die Knoten werden jeweils gleichzeitig (*parallel*) aktualisiert. Die Abfolge der Aktualisierungen definiert eine Zeitachse im System. Wegen der endlichen Anzahl aller möglichen Netzwerkzustände und der deterministischen Zeitentwicklung ist die zeitliche Dynamik nach einer Übergangszeit periodisch. In der Übergangszeit werden *transiente* Netzwerkzustände durchlaufen. Danach befindet sich das Netzwerk auf einem *Attraktor* der Dynamik bzw. auf einer periodischen Bahn, wenn es keine transienten Zustände gibt.

Wenn man die Dynamik von einigen Netzwerkrealisierungen in einer Simulation auf einem Computer verfolgt, dann stellt man fest, dass auf dem jeweiligen Attraktor nur ein Teil aller Knoten ihren Zustand ändert (*blinkt*). Wenn man in diesen Prozess eingreift, den Zustand eines zufällig ausgewählten Knotens umkehrt und die Dynamik dann weiter laufen lässt, dann wirkt sich das nach einer Übergangszeit auf eine Untermenge der blinkenden Knoten aus, zusätzlich können vorher *stabile* Knoten anfangen zu blinken und blinkende Knoten können stabilisiert werden. Die Gesamtzahl der Knoten, die ihr Verhalten bei einem lokalen Eingreifen ändern, kommt immer als eine “Überraschung”: manchmal ist sie klein oder 0, manchmal ändern viele Knoten ihr Verhalten. Die Antwort des Systems ist vielfältig, gleichzeitig aber auch strukturiert. Man kann nämlich statistische Vorhersagen für verschiedene Systemantworten angeben. Das System ist *komplex*! Eine abstrakte Definition der Komplexität folgt in Kürze, hier ist noch anzumerken, dass diese Komplexität im zeitlichen Verhalten sich auf eine “duale” Weise auch in den “räumlichen” Eigenschaften im Konfigurationsraum der Netzwerke widerspiegelt. Man findet in verschiedenen Netzwerkrealisierungen im Allgemeinen verschiedene auf allen Attraktoren stabile Knotenuntermengen, gleichzeitig kann man für jedes feste N die Anzahl der auf allen Attraktoren stabilen Knoten statistisch vorhersagen.

Wie am Anfang erwähnt, kann man für Modelle komplexer Systeme zwei Arten von Fragestellungen unterscheiden. Zum Einen untersucht man die Gemeinsamkeiten in verschiedenen komplexen Systemen, indem man einzelne Modelle komplexer Systeme im Detail studiert um die Komplexität mathematisch bzw. anschaulich nachvollziehbar zu machen. Wenn verschiedene Systeme sich aufeinander abbilden lassen, dann spricht man von einer *Universalität* im Verhalten. Die Existenz von universellen Mechanismen und Strukturen, die komplexen Systemen innewohnen,

hat sie für die Physik erst interessant gemacht. Die Existenz komplexer Systeme mit Vielfalt und Ordnung ist für sich genommen eine universelle Eigenschaft. Stuart Kauffman hat mit seinem Modell der booleschen Zufallsnetzwerke einen wichtigen Beitrag zum Aufzeigen dieser Universalität und der Wege zu ihrem Studium und in der Organisation der wissenschaftlichen Bemühungen bei der Untersuchung komplexer Systeme geleistet.

Per Bak mit seinen Kollegen hat vermutlich als Erster einen universellen Mechanismus der Entstehung von komplexem Verhalten beschrieben: die *selbstorganisierte Kritikalität*. Welchen Stellenwert komplexes Verhalten in unserem Weltbild einnimmt und, insbesondere, wie die allgegenwärtige Komplexität aus physikalischen Gesetzen entsteht, warum die Komplexität *emergent* (*auftauchend*) ist, umreißt die zweite Hauptrichtung der Forschung über komplexe Systeme.

Ich werde mich hauptsächlich mit dem Verständnis der booleschen Zufallsnetzwerke auseinandersetzen, insbesondere mit der Universalität in den entsprechenden Modellen, und am Rande die Schnittstellen zu der selbstorganisierten Kritikalität diskutieren. Als Nächstes betone ich die Eigenschaften der booleschen Zufallsnetzwerke, die sie gemeinsam mit vielen anderen komplexen Systemen haben, und stelle relevante Begriffe vor.

Wie die meisten betrachteten Modelle komplexer Systeme bestehen die booleschen Zufallsnetzwerke aus vielen Elementen (Knoten), die lokal (mittels der Verknüpfungen) einander beeinflussen bzw. ein zunächst lokal korreliertes Verhalten zeigen. (Die lokalen Korrelationen werden durch die Netzwerk - Topologie und die Wahl der Funktionen definiert.) Dass dabei *übergeordnete globale Strukturen* (die oben angesprochenen statistischen Gesetzmäßigkeiten) entstehen, kann treffend durch den Satz: “Das Ganze ist mehr als die Summe seiner Teile”, ausgedrückt werden. Eine ausführlichere Betrachtung folgt im Abschnitt über die Komplexität.

Die globalen dynamischen Strukturen lassen auf die Existenz (auftauchender) langreichweitiger Korrelationen schließen. Aus der Analogie mit der Physik der Phasenübergänge in Gleichgewichtssystemen erscheint plausibel, dass die globalen Strukturen häufig mit dem Auftreten von *Potenzgesetzen* einhergehen. Andererseits wird das Auftreten von Potenzgesetzen per Definition als komplexes Verhalten angesehen, weil die Potenzgesetze *Selbstähnlichkeit* über mehrere Dekaden im System mit sich bringen. Wenn man einen Ausschnitt des Systems vergrößert, ob “zeitlich” oder “räumlich”, findet man dieselben Gesetzmäßigkeiten, keine Zeit und keine Länge ist charakteristisch bzw. ausgezeichnet. Für das Standardmodell der booleschen Zufallsnetzwerke findet man, dass die Wahrscheinlichkeiten in den statistischen Gesetzmäßigkeiten (wie die Wahrscheinlichkeit dafür, einen gegebenen Anteil der auf den Attraktoren blinkenden Knoten zu finden) mit der Systemgröße N skalieren. Erst in den Jahren 2002 bis 2005 ist es, auch dank dieser Arbeit, gelungen, die richtige Skalierung im beschriebenen Modell boolescher Zufallsnetzwerke zu finden, vgl.

1. Einleitung

Abschn. 2.4. Potenzgesetze und Selbstähnlichkeit sind ein häufiges Wahrzeichen der Komplexität⁴.

Die Systeme, die durch Potenzgesetze und Skalenverhalten charakterisiert werden, werden *kritisch* genannt. Diese Bezeichnung kann, analog zur Theorie der Phasenübergänge⁵, darauf hinweisen, dass das komplexe Verhalten sich bei einer Feinabstimmung von Modellparametern ergibt und verschiedene Verhaltensmuster (*Phasen*) im System trennt. In booleschen Zufallsnetzwerken ist die Anzahl der eingehenden Verbindungen K je Knoten ein solches Parameter, falls alle booleschen Funktionen mit derselben Wahrscheinlichkeit gewählt werden. Nur bei $K = 2$ findet man ein komplexes Verhalten. Bei kleinerem K (trivialerweise bei $K = 0$) sind in einem Netzwerk aus dem Netzwerk - Ensemble im Mittel alle Knoten stabil auf allen Attraktoren, man spricht von der *geordneten Phase*. Bei $K > 2$ blinkt in großen Netzwerken durchschnittlich ein nicht verschwindender Anteil aller Knoten, typische Attraktoren sind von der Größe des gesamten Zustandsraums, das System befindet sich in der *chaotischen Phase*, vgl. Abschn. 2.4. Wie auch im Allgemeinen gültig, befindet sich die Phase mit dem komplexen Verhalten im Parameter - Phasenraum an der Grenze zur Phase mit dem chaotischen Verhalten. Bei der weiter unten beschriebenen selbstorganisierten Kritikalität sieht man in der komplexen Phase selbst katastrophale Vorgänge, die mit einer kleinen Wahrscheinlichkeit auftreten und das gesamte System betreffen. Man spricht von der Komplexität “an der Grenze zum Chaos” (“at the edge of chaos”⁶).

Zufall spielt eine wesentliche Rolle für das Auftauchen der Komplexität. Der Zufall im Sinne von Gleichverteilung auf einer bestimmten Ereignismenge (zufällige Zuweisung der Eingänge und der Funktionen in booleschen Zufallsnetzwerken, weißes Rauschen im weiter unten beschriebenen Modell der selbstorganisierten Kritikalität) ist plausibel und daher in physikalischen Modellen als ein Ausgangspunkt sehr beliebt⁷. Jede Form von Struktur bedarf dagegen einer Erklärung. Und der Zufall zusammen mit lokal definierten Korrelationen kann zu globalen Korrelationen, insbesondere zur Komplexität im Modellsystem führen. Probabilistische Definition von booleschen Zufallsnetzwerken weist Parallelen auf zu ungerichteten Zufalls-

⁴Im Gegensatz zu den Fraktalen in der Mathematik findet man Selbstähnlichkeit in komplexen Systemen ausschließlich im statistischen Sinn und nicht für einzelne Systemrealisierungen. Mit der Forderung nach vielen Systemrealisierungen für ein komplexes System (siehe Abschn. 1.2) ist kein Fraktal für sich genommen komplex.

⁵Beispielsweise wird bei der kritischen Temperatur kritische Opaleszenz zwischen der flüssigen und gasförmigen Phase beobachtet.

⁶S. Kauffman hat genauer den Begriff “life at the edge of chaos” geprägt und zwar im Zusammenhang mit den genetischen regulatorischen Netzwerken, deren komplexe Regulation flexibel ist, spricht, durch Mutationen im Laufe der Evolution zu einem großen Teil umorganisiert werden kann. Er hat vermutet, dass diese Umorganisation eine intrinsische, die Darwin’sche natürliche Selektion ergänzende Eigenschaft des Systems ist.

⁷In einer weiter führenden Diskussion könnte man den Zufall, also (gleich)wahrscheinliches Vorkommen von Ereignissen aus einer Ereignismenge, physikalisch aus dem zweiten Hauptsatz der Thermodynamik begründen, dessen eigene Herkunft umstritten ist. In jedem Fall sind zufällige Prozesse allgegenwärtig, angefangen von Vakuumfluktuationen und bis hin zu komplexen Systemen, die Zufall produzieren, wie die Farbpartikel - Verteilung beim Streichen der Wände.

graphen, wo jedes Knotenpaar mit einer Wahrscheinlichkeit p verbunden ist. Die kritischen Eigenschaften der Zufallsgraphen sind seit den 1960er Jahren bekannt, erstaunlicherweise hat das nicht zu einem früheren Durchbruch im Verständnis der booleschen Zufallsnetzwerke geführt. Die Zufallsgraphen gelten als eines der besten Beispiele der Mächtigkeit probabilistischer Methoden, mit deren Hilfe z.B. gezeigt wurde, dass ab einem kritischen Wert von p sprunghaft globale Strukturen entstehen. Beim Ausnutzen der Parallelen muss man vorsichtig sein, da boolesche Zufallsnetzwerke gerichtet sind und außerdem ein dynamisches System darstellen. Auch bestehende Parallelen zu Gitter - Perkolationsproblemen sind beschränkt nützlich.

Die Methoden, die auf boolesche Zufallsnetzwerke bislang angewendet wurden, basieren auf den aus der Mathematik und Physik bekannten probabilistischen Ansätzen. Wegen der nicht trivialen Natur des Problems blieb deren konsequente Anwendung jedoch bis vor kurzem aus. In der Liste der angesprochenen Methoden wären hier vor allem die Theorie der *stochastischen Prozesse* aus der statistischen Physik, die Stochastik (die Wahrscheinlichkeitstheorie und die mathematische Statistik), die Kombinatorik, die Theorie der Phasenübergänge, insbesondere Skalierungs- und Renormierungsansätze, die Theorie der nichtlinearen dynamischen Systeme, die Chaostheorie und die Logik zu nennen. Allgemein finden im Rahmen der Wissenschaft über Komplexität vereinzelt auch "exotische" Herangehensweisen wie die Feldtheorien aus der Elementarteilchenphysik Anwendung. In den jeweiligen Modellen aus der Chemie, Physik, Biologie, Erdkunde, Soziologie, uvm. sind auch entsprechende Konzepte und Methoden aus diesen Wissenschaften wichtig. Ich werde beispielsweise die Relevanz der Reaktionskinetik im Zusammenhang mit der Beschreibung der genetischen regulatorischen Netzwerke diskutieren.

Numerische Simulationen werden üblicherweise zur Ergänzung und Bestätigung der analytisch gewonnenen Erkenntnisse herangezogen, wobei diese Simulationen häufig an ihre Grenzen stoßen. In booleschen Zufallsnetzwerken treten die interessanten *asymptotischen* Eigenschaften erst bei großen N auf, wo *numerische Simulation eingeschränkt anwendbar* sind. Im Hauptteil dieser Arbeit werde ich einige Male darauf eingehen.

Das oben definierte einfache Standardmodell der booleschen Zufallsnetzwerke kann auf verschiedene Weisen erweitert werden, siehe das Kapitel 2. Unter anderem werden intensiv Erweiterungen im Hinblick auf die realistischere Beschreibung von genetischen regulatorischen Netzwerken studiert. Mir ist an dieser Stelle wichtig zu erwähnen, dass diese Beschreibung eher als ein Nebenprodukt der Forschung betrachtet werden sollte und das Verständnis der Dynamik eines wichtigen generischen komplexen Modells wie die booleschen Zufallsnetzwerke für sich genommen wertvoll ist.

Boolesche Zufallsnetzwerke stellen ein Modell dar, das keine selbstorganisierte Kritikalität (*SOC*) zeigt. Ohne weitgehende Erweiterungen des Modells (*NKCS* - Modell, siehe Abschn. 1.3) muss man ein Parameter ($K = 2$) genau einstellen um komplexes Verhalten zu sehen. In *SOC* - Systemen dagegen ist der kritische Zustand in einem gewissen Sinne (vgl. Abschn. 1.3 über *SOC*) ein Attraktor der

1. Einleitung

Dynamik. In der wissenschaftlichen Umgangssprache verwendet man hierfür die Metapher “Ordnung gratis” (“order for free”).

Das Standardbeispiel von einem SOC - System ist ein Sandhaufen. Man lässt Sand von oben auf eine Platte rieseln, am Rand kann der Sand entweichen. Nach einer gewissen Zeit erreicht die Steigung des Sandhaufens einen kritischen Wert. Ein Sandkorn kann nun wegen verringerter Reibung eine große *Lawine* auslösen, indem es beim Abrollen nach unten weitere Sandkörner in Bewegung setzt. Größere Lawinen, die einen endlichen Anteil von Sandkörnern in Bewegung setzen repräsentieren *katastrophale Ereignisse* und kommen selten aber nicht periodisch vor, alle kleineren Lawinengrößen werden beobachtet. Das System ist komplex. Die Lawinengrößenverteilung ist ein Potenzgesetz mit einem negativen Exponenten. In der Natur werden größere Lawinen besser durch Potenzgesetze beschrieben, bei kleineren Lawinen sind die nicht universellen Eigenschaften des konkreten Systems wichtig. In einem idealisierten Modell wird Zufall durch Herablassen der Sandkörner an einer beliebigen Position auf der Platte eingeführt. Man spricht auch vom langsamen Treiben, um die zeitliche *Skalentrennung* im System hervorzuheben, die Lawinen stellen nämlich die schnelle Komponente der Dynamik dar. Außerdem sind die SOC - Systeme wie auch die meisten komplexen Systeme *offene Nichtgleichgewichtssysteme*, charakterisiert durch Energie- oder Teilchenfluss durch das System. SOC - Systeme sind *dissipativ*. Wenn die Energie oder Teilchen das System am Systemrand verlassen, spricht man von *absorbierenden Randbedingungen*.

Im Standardmodell der booleschen Zufallsnetzwerke könnte man *Mutationen*, wie zum Beispiel zufällige lokale Veränderungen eines Knotenzustandes, als langsames Treiben interpretieren und das anschließende Laufen auf den Attraktor als eine Lawine, die Lawinengrößen - Verteilung ist dann ein Potenzgesetz mit dem Exponenten $-3/2$.

Zusammenfassend sind boolesche Zufallsnetzwerke ein komplexes System. Nach dieser Behauptung möchte ich den umfassenden Begriff der Komplexität etwas ausführlicher betrachten.

1.1. Beispiele komplexer Systeme

Es geht um komplexe Systeme, wenn man darüber nachdenkt, wie aus einem Embryo ein erwachsener Organismus entsteht, wo die Verbindung zwischen dem Wasser und der Pracht der Eisblumen liegt, wenn man sich die Frage stellt, wie Kulturen kommen und gehen, wie Börsenkurse schwanken, wie Menschen ihre Meinung bilden, warum jeder einen Bekannten hat, der Kofi Annan persönlich kennt, woraus die Erfolgsgeschichte des Internets resultiert, wie sich Ökosysteme bilden, wie Kohlenoxid - Ablagerungen die Effizienz der Brennstoffzellen erhöhen können, wie man Verkehrsstaus meidet, wie neuronale Netzwerke im Gehirn Muster erkennen, wie Sanddünen charakteristische Formen annehmen, wie sich Waldbrände ausbreiten, wie Lawinen in den Bergen abgehen und Erdbeben unvorhergesehen kommen, wie Tierfellmuster entstehen, wann Kriege ausbrechen und wie lange sie dauern, mit

welcher Intensität die Quasare aufleuchten, und wie aus dem Chaos der Ursuppe alles Leben so unglaublich schnell hervorging.

Die angesprochenen Phänomene und Systeme werden seit einigen Jahrzehnten mit viel Inspiration untersucht. Wir fangen an zu verstehen, wie das diffuse Uhrwerk der Entstehung der uns umgebenden Schönheit funktioniert.

Anfang der 1990er Jahre hat Alan Perelson mit einem einfachen Räuber - Beute Modell anhand von Daten aus dynamische - Antwort - Studien bestimmter Medikamente gezeigt, dass in der Latenzperiode HIV Viren sehr schnell repliziert werden und daher auch sehr schnell vom Immunsystem wieder vernichtet werden müssten. Diese Erkenntnis führte zur Entwicklung von kombinierten Therapien für HIV - Infizierte im frühen Stadium [Pere 96], wobei eine Kombination von Medikamenten das Immunsystem schont. Viele andere Beispiele der Anwendung der Erkenntnisse über komplexe Systeme wie im Militär oder im Bildungssystem findet man im neuen Buch des Physikers Yaneer Bar Yam "Making Things Work" [Bar 03].

Da ein komplexes System komplex ist, ist es schwer, ein "einfaches" komplexes System zu finden. Notwendigerweise wird ein solches System minimal komplex sein. Ich erinnere mich an dieser Stelle an ein jedem Kind vertrautes Spielzeug: das Kaleidoskop. Das Wort "Kaleidoskop" kommt aus dem Griechischen. Wörtlich übersetzt bedeutet es "Schönbildseher" oder "Schönbildschauer". Bunte wechselnde Muster, die man bewundern kann, indem man das Rohr des Kaleidoskops um seine Achse dreht, sind komplexe Abbilder von einigen wenigen bunten Plastikteilchen im durchsichtigen Döschen am Ende des Rohres. Drei längs des Rohres in Form eines Dreiecks aufgestellte Spiegel führen zum im Prinzip unendlichfachen Spiegeln der Plastikteilchen, wobei die Symmetrie der Spiegelaufstellung im Bild sich wieder findet. Die Spiegel stellen eine große Menge von dicht aneinander liegenden Bildelementen im beschränkten Sichtfeld her, ihre Anordnung rührt von der einfachen "lokalen" Anordnung der Plastikteilchen her, die symmetrische Aufstellung der Spiegel bringt Struktur in jedes entstehende Bild, die vielfache Spiegelung führt zu anschaulich komplexen Mustern.

Einige der vielen mit physikalischen Methoden untersuchten Modelle mit komplexem Verhalten sind klassische kritische Phänomene wie die kritische Opaleszenz mit den nach einem Potenzgesetz verteilten Fluktuationsgrößen; Gitter - Perkulationsprobleme, bei denen sich gleichzeitig mit der Ausbildung der größten Komponente die Größen der zusammenhängenden Gebiete nach einem einfachen statistischen Gesetz verteilen; Wachstumsprozesse auf Oberflächen, die komplexe emergente raue Oberflächenstrukturen produzieren können, u.a. die der Schneeflocken; Modelle der Proteinfaltung; Spingläser, die z.B. durch eine zum Heisenbergmodell ähnliche Hamilton - Funktion beschrieben werden und viele metastabile Konfigurationen aufweisen; Meinungsbildungsmodelle von kommunizierenden Individuen mit konkurrierenden Meinungen (mit *Frustration*), hydro- und aerodynamische Modelle turbulenter Strömungen, die in einem bestimmten Parameterbereich komplexe Strukturen wie Bénard - Konvektionsrollen ausbilden; Reaktion - Diffusion - Systeme wie die Belousov - Zhabotinsky Reaktion mit zeitlich und räumlich in chemischen Reaktionen aufgrund von oszillierender Konzentration von Ce^{4+} variierenden

1. Einleitung

farbigen Mustern; zelluläre Automaten wie das “game of life”⁸; SOC - Modelle (siehe Abschn. 1.3); verschiedene Variationen von booleschen Zufallsnetzwerken und Netzwerken im Allgemeinen und andere. Die meisten dieser Modelle wurden letztendlich durch Untersuchungen realer komplexer Vorgänge wie der Vorgänge hinter den Beispielen am Anfang des Abschnitts motiviert.

Es folgen Definitionen und ergänzende Informationen, die einer genaueren Umschreibung des Kontexts dienen, in dem die komplexen Systeme wie die booleschen Zufallsnetzwerke abgegrenzt werden.

1.2. Komplexität

Ich muss hier die Begriffe *das komplexe System* und *die Theorie der Komplexität* genauer definieren und abgrenzen, weil aufgrund der umfassenden und auf sich selbst bezogenen (autologischen) Bedeutung des Wortes komplex⁹ die Begriffe je nach Wissenschaftsgebiet und je nach Autor unterschiedlich verwendet werden¹⁰. Insbesondere wähle ich die Definitionen, die meiner Meinung nach in vielen Wissenschaftsgebieten im Wesentlichen ihre Gültigkeit beibehalten, während eine Überschneidung der Definitionsbereiche der verwandten Begriffe wie Selbstorganisation, Emergenz, selbstorganisierte Kritikalität und Komplexität ausgeschlossen wird.

Zunächst meine ich unter der Theorie der Komplexität weder die Rechenkomplexität in der theoretischen Informatik, wo man Probleme nach erforderlichen Ressourcen zu ihrer Lösung klassifiziert, noch das strategische Management in den Wirtschaftswissenschaften, wo das Zusammenspiel von verschiedenen Strategien von vielen Firmen untersucht wird. Diese Beispiele können zwar im Rahmen der folgenden Definitionen erfasst werden, befassen sich jedoch hauptsächlich mit der Lösung eines konkreten Problems und nicht mit dem universellen Verständnis.

Auch die chaotischen Systeme und die Gleichgewichtssysteme, wenn deren statistische Beschreibung keine vielfältigen zeitlichen und/oder räumlichen Strukturen mit langreichweitigen Korrelationen aufweist, werden als nicht komplex angesehen. Komplex ist auch nicht einfach zahlreich. Die algorithmische Komplexität im Sinne von der minimalen Information (Zeit), die zur Beschreibung (Erzeugung) des Systems benötigt wird, bemühe ich auch nicht.

⁸In “game of life” kann jede Zelle auf einem zweidimensionalen Gitter zwei Zustände haben, tot oder lebendig, sie stirbt wenn sie nicht genau 2 oder 3 (aus 8 möglichen) lebendigen Nachbarn hat und wird geboren, wenn sie genau 3 Nachbarn hat. Diese und keine anderen Regeln mit 2 Zuständen auf einem zweidimensionalen Gitter führen zur Entstehung, je nach Anfangszustand, von komplexen Mustern.

⁹Lateinisches Wort “complectari” (umfassen) bzw. “plectere” (flechten) wurzelt im griechischen “pleko” oder “plektos” (verweben).

¹⁰Man kann Komplexität im Hinblick auf die Elemente, auf die Beziehungen, auf die Beschreibbarkeit, die Stabilität und dergleichen definieren. Bekannte Autoren wie Stafford Beer, William Ashby und viele andere haben in diesem Kontext Worte wie Varietät, Kompliziertheit, Verschiedenartigkeit, Konnektivität, Robustheit, Kapazität, Intensität etc. verwendet.

Ein System bringt die vom Beobachter identifizierbaren Teile in einen Zusammenhang. Ich nenne ein aus vielen Teilen (Systemkomponenten) bestehendes dynamisches System, in dem Wechselwirkungen lokal definiert sind, *korreliert*. Ein korreliertes System, das ein vielfältiges und strukturiertes Verhalten zeigt, nenne ich ein *komplexes System*. Vielfältig meint, dass das Verhalten vom System qualitativ reichhaltiger ist als dasjenige von einzelnen Komponenten unter Berücksichtigung ihrer lokalen Wechselwirkungen. Das Wort qualitativ wird nicht näher spezifiziert, auf jeden Fall wird aber damit gemeint, dass das Gesamtverhalten keine Superposition von den Verhalten von einzelnen Komponenten ist: “Das Ganze ist mehr als die Summe seiner Teile”¹¹. Strukturiert meint, dass es probabilistische Gesetzmäßigkeiten gibt, die den Phasenraum des Systems strukturieren lassen; dadurch wird die Vielfalt eingeschränkt. Die globalen Strukturen lassen auf die Existenz von einem kooperativen Verhalten von vielen bis von allen Systemkomponenten schließen, wobei viele verschiedene Kooperationen (sie werden häufig als *metastabile Zustände* bezeichnet) entstehen können. Nennenswert ist, dass die angegebenen Definitionen präziser sind im Vergleich zur Definition komplex sei vielfältig, die beispielsweise von Per Bak benutzt wurde.

Die geeignete mikroskopische Beschreibung der anschaulich komplexen Naturphänomene beinhaltet häufig kontinuierliche Gleichungssysteme. Die angegebene Definition der Komplexität bezieht sich im Normalfall jedoch auf diskrete Systemkomponenten. Damit wird eine praktische Vereinfachung der Modelle und die mit der Suche nach Universalität in komplexen Systemen zusammenhängende Reduktion auf das Wesentliche beabsichtigt. Außerdem bezieht sich die Definition offensichtlich auf ein klassisches System, die Komplexität in quantenmechanischen Systemen ist bislang kaum angegangen worden, eine Ausnahme ist die quantenmechanische Betrachtung der Systeme fern vom Gleichgewicht [Anta 97].

Verschiedene Definitionen der Komplexität von anderen Autoren [Gold 99a] heben verschiedene zusammenhängende Merkmale bekannter komplexer Systeme hervor wie variierende globale Strukturierung (Goldenfeld und Kadanoff), Abwesenheit von charakteristischen Zeit- oder Raumskalen, Sensitivität bezüglich des Anfangszustands bzw. Existenz von Attraktoren und vielen auseinander laufenden Trajektorien im Konfigurationsraum (Whitesides und Ismagilov), dauernde kaum vorher-sagbare Änderung des komplexen Systems (W. Brian Arthur), heterogene Bestandteile mit heterogenen Wechselwirkungen (D. Rind); das Verhalten eines komplexen Systems ist schwer bis unmöglich deterministisch auf einer höheren Abstraktionsebene zu verstehen, nachzubauen oder vorherzusagen (Weng, Bhalla und Iyengar).

Eine Untersuchung der systemtheoretischen Aspekte der Komplexität ist wichtig zum Verständnis der modellierten Systeme. Insbesondere werde ich in dieser Arbeit im Kontext der Beschreibung von echten genetischen Netzwerken die Zu-

¹¹In der Philosophie ist diese Aussage auch gleichzeitig die Grundidee des Holismus (vom griechischen “holon” - Einheit), wobei so genannter gieriger Reduktionismus den Gegensatz dazu bildet.

1. Einleitung

sammenhänge zwischen dem Aufbau der Modelle boolescher Zufallsnetzwerke und der Funktion der entsprechenden genetischen Netzwerke beschreiben.

Eine Betrachtung der historischen Entwicklung der Theorie der Komplexität beleuchtet die möglichen Weiterentwicklungen und die Austauschmöglichkeiten mit verwandten wissenschaftlichen Disziplinen. Ich halte diesen Austausch für wichtig und beschreibe die Entwicklung der Theorie der Komplexität im Anhang A.

Die Eigenschaften komplexer Systeme, die im Zusammenhang mit Mechanismen der Entstehung der Komplexität interessant sind, werden nun am berühmten Beispiel der selbstorganisierten Kritikalität erläutert.

1.3. Selbstorganisierte Kritikalität

Ein dissipatives korreliertes System bezeichne ich *potentiell selbstorganisiert kritisch*, wenn es auf einer langsameren als die Dynamik des Systems Zeitskala auf eine zufällige Weise getrieben wird, insbesondere durch lokale Veränderungen durch einen Zufallsprozess. Dissipativ meint, dass die Entropie des Systems kleiner werden kann. Beispielsweise kann die Dissipation durch absorbierende Randbedingungen realisiert werden. Da die Dissipation mit der Dynamik im System zusammenhängt, spricht man von der *schnellen Dissipation*. Die erwähnten Zufallsprozesse werden als *langsames Treiben* bezeichnet. Das langsame Treiben und die schnelle Dissipation weisen auf die Offenheit des Systems und auf die Existenz metastabiler Zustände hin. Ein *selbstorganisiert kritisches System (SOC)* ist ein potentiell selbstorganisiert kritisches System, das sich ohne Eingreifen von außen zu einem robusten komplexen System entwickelt, dessen Strukturen durch Potenzgesetze charakterisiert werden. *Robust* meint, dass die Komplexität nicht durch kleine Änderungen von Parametern zerstört werden kann. Das Auftauchen von Potenzgesetzen wird durch das Wort "kritisch" betont. Man vermutet, dass viele potentiell selbstorganisiert kritische Systeme auch selbstorganisiert kritisch sind. Selbstorganisiert kritische Systeme sind also offene Systeme mit zufallgetriebener Zeitskalentrennung - diese Systeme werden langsam durch einen Zufallsprozess (*Fluktuationen, Rauschen, Mutationen etc.*) getrieben und kommen schnell in einen Zustand niedrigerer Entropie bzw. dissipieren schnell die Energie -, die ein komplexes kritisches - charakterisiert durch Auftreten von Potenzgesetzen - Verhalten als einen Attraktor der Dynamik haben. Offensichtlich ist der Übergang zum komplexen Verhalten selbst komplex.

Die Motivation für die Betrachtung der selbstorganisierten Kritikalität ist analog zu derjenigen der Modelle wie das Ising - Modell in der Gleichgewichtstatistik, die die wesentlichen Elemente der realen Systeme beinhalten, die Symmetrie und die Austausch - Wechselwirkung im Falle des Ising - Modells. Man sucht nach universellen Mechanismen der emergenten Komplexität. Aus diesem Grund erscheinen SOC - Modelle abstrakt im Vergleich zu (realen) Systemen, die sie beschreiben.

Kritisches Verhalten wird in der SOC - Definition als Synonym für Potenzgesetze verwendet. Wegen der Randbedingungen existiert *Skalenverhalten* (Parameterreskalierung - unabhängige funktionale Abhängigkeit der Observablen) nicht zwin-

gend in SOC - Systemen. Im Kontext der booleschen Zufallsnetzwerke werde ich später mit dem Wort “kritisch” die komplexe Phase an der Grenze zu Chaos bezeichnen. Allgemein bezeichnet man bekanntlich das Verhalten am Phasenübergang als kritisch (*kritische Exponenten* wie die Exponenten der Potenzgesetze bei SOC, kritisches Skalenverhalten, kritische Parameterwerte), sowohl im Gleichgewicht als auch im Nichtgleichgewicht. In der Literatur wird das Wort “kritisch” vereinzelt auch anders verstanden: Zum Beispiel hat Newman im Zusammenhang mit seinem evolutionären Modell vom Aussterben von Spezies unter kritisch die Existenz koevolutionärer Anpassungs - Lawinen verstanden.¹²

Beispiele für (ggf. idealisiert) selbstorganisiert kritische Systeme sind [Newm 99] “Sandhaufen” aus länglichen Reiskörnern, diskrete Sandhaufenmodelle mit einem lokal definierten Schwellenwert für die Steilheit des Sandhaufens und mit Sand - Dissipation an Nachbarplätze, vermutlich einige Systeme mit dem so genannten $1/f$ - Rauschen¹³, das boolesche NKCS - Koevolution - Modell¹⁴ mit fluktuierendem K und selektivem Wettbewerb von Kauffman und Johnsen, diskrete Waldbrandmodelle, Modelle von Massensterben von Spezies im Laufe der Evolution (Bak - Sneppen Modell, Solé - Manrubia Modell, das Modell von Newman), Olami - Feder - Christensen Erdbebenmodell, Ausbildung von Staus, Ausfälle in Elektrizitäts - Netzwerken während einzelne Verbraucher nur kleine zufällig variierende Stromleis-

¹²Obwohl Newman ein Modell ohne explizite koevolutionäre Lawinen gefunden hat und sich dadurch von einer Betrachtung der Evolution als einem selbstorganisiert kritischen System im Sinne von Bak, Sneppen und anderen distanziert hat, zeigt sein Modell qualitativ die Potenzgesetze, die man aus experimentellen Daten kennt und genügt auch sonst der hier angegebenen Definition eines selbstorganisiert kritischen Systems. In wenigen Worten, auch um ein weiteres SOC - Beispiel zu geben, betrachtet Newman N Spezies mit zufällig zugewiesenen reellen Fitnesswerten aus dem Intervall von 0 bis 1. In jedem ungeraden diskreten Zeitschritt, dem Evolutionsschritt, wird ein kleiner Anteil f der Spezies “selektiert”, indem sie neue zufällige Fitnesswerte bekommen. In dem Anpassungsschritt an die zufällige feindliche Umgebung $\eta(t) \in [0, 1]$ zu den geraden Zeiten t sterben alle Spezies, deren Fitnesswerte kleiner als $\eta(t)$ sind, und werden durch neue Spezies mit zufälligen Fitnesswerten ersetzt. Die Spezies betrachte ich als Energiepakete. Das Ausradieren von “benachbarten” Spezies mit kleinen Fitnesswerten realisiert die Korrelationen im System und sorgt gleichzeitig für das langsame Treiben im System, die Fitnesswerte werden dadurch langfristig größer. Die Zuweisung der neuen Fitnesswerte im Allgemeinen stellt die Dynamik im System dar und wirkt dem Anwachsen der Fitnesswerte entgegen (die schnelle “Dissipation”). Eventuell kommt es zu Massensterben.

¹³In Systemen mit $1/f$ - Rauschen fällt die spektrale Energiedichte des Frequenzspektrums bestimmter Signale oder Prozesse umgekehrt proportional zur Frequenz ab. Ursprünglich gefunden um 1925 in der Triode in Elektronik wurde das Anwendungsgebiet des Begriffes im Laufe der Jahre erweitert auf (ggf. missbraucht für) verschiedenartige Prozesse. Man kann die Gleichgewichtsprozesse in einer Triode nicht als selbstorganisiert kritisch bezeichnen, wenn auch die Beschreibung des entsprechenden Phänomens durch Überlagerung von gaußschen Relaxationsprozessen Parallelen zu SOC aufweisen mag. Selbstorganisierte Kritikalität kann jedoch in der zeitlichen Variation des Lichtintensität - Signals von einem Quasar, im Gutenberg - Richter - Gesetz für Erdbeben, im Zipf - Gesetz für die Verteilung der Größe der städtischen Bevölkerung und anderen vermutet werden, die als Systeme mit $1/f$ - Rauschen betrachtet werden.

¹⁴Fitnessbeitrag von einem Gen wird von je C Genen von S Spezies beeinflusst, die auf NK Fitness - Landschaften existieren.

1. Einleitung

tungen benötigen, Wachstum von Unternehmen als eine Spiegelung der variierenden Nachfrage, (bestimmte Auswertungen von) Intensitäten von Kriegen, vielleicht sogar die Ausbildung der Galaxien und der Planetensysteme in der Astrophysik. Viele dieser Modelle hat man bislang nur wenig verstanden.

Wenngleich die Idee der universellen selbstorganisierten Kritikalität viele Wissenschaftler fasziniert, so ist ihre Existenzberechtigung noch lange nicht sicher, abgesehen von ihrer unbestreitbaren Rolle in der Bündelung der Forschungsaktivitäten. Die Kontrahenten der SOC stellen ihre Bedeutung bei der Erklärung der Komplexität in Frage, sie argumentieren, dass ungeachtet der Entdeckung von weiteren Beispielen von SOC und der Anwendung immer anspruchsvollerer Methoden man nicht einmal im Falle des Sandhaufenmodells von Bak, Tang und Wiesenfeld zu einem tieferen Verständnis gelangt ist. Auch eine breite Anwendbarkeit der Erkenntnisse ist fraglich. Experimente mit realen Sandhaufen und anderen granularen Systemen [Nage 92, Dalt 01] und auch analytische Überlegungen [Jens 89] haben für die Verteilung der Lawinengrößen ein Potenzgesetz $1/f^2$ gezeigt, und das nach einem Plateau bei kleineren Lawinengrößen bzw. kleineren Mengen der dissipierten Energie. Gleichzeitig hat das Energiespektrum der brownischen Bewegung in einer Dimension trivialerweise die Frequenzabhängigkeit $1/f^2$. Experimente mit Sandhaufen mit runden Reiskörnern haben gestreckte exponentielle Lawinengrößenverteilungen ergeben, wobei die Trägheit der größeren Lawinen in diesem Fall eine wesentliche Rolle spielt. Bei Erdbeben, Kriegen und Aussterben von Spezies sind auch andere als Potenzgesetze Abhängigkeiten mit den experimentellen Daten vereinbar, bzw. Potenzgesetze sind nur über wenige Dekaden auf einer logarithmischen Skala zu beobachten. Es ist zu bezweifeln, dass bei diesen Phänomenen derselbe Mechanismus über mehrere Dekaden in Raum und Zeit wirkt. Auch ist die Selbstorganisation an sich in SOC-Modellen fraglich, da das langsame Treiben bedeutet, dass die treibende Kraft eine globale Kenntnis über das System besitzt, nämlich, wann die verursachte dissipative Dynamik zu Ende ist [Dick 02]. In jedem Fall findet man bei potentiell selbstorganisiert kritischen Systemen eine sich über mehrere Dekaden erstreckende Größenverteilung der Ereignisse, die nach Definition komplex ist.

2. Überblick: Boolesche Zufallsnetzwerke (RBN)

2.1. Einführung

Im einleitenden Kapitel 1 habe ich das 1969 von S. Kauffman ursprünglich zur Modellierung der genetischen regulatorischen Netzwerke eingeführte Modell [Kauf 69] beschrieben. Dieses Netzwerkensemble aus N zufällig verknüpften Knoten mit booleschen Werten, $K = 2$ Eingängen je Knoten und 16 Knotenfunktionen, die alle gleichwahrscheinlich sind, bezeichne ich im Folgenden als gewöhnliche boolesche Zufallsnetzwerke oder das Standardmodell oder *sRBN* (standard random Boolean networks). Die Knotenverbindungen im sRBN - Modell sind gerichtet, die Aktualisierung erfolgt synchron für alle Knoten. Das Modell ist diskret, dynamisch und dissipativ in dem Sinne, dass in einem Netzwerkzustand die Information über die Phasenraumtrajektorie ausstrahlt wird, wenn er aus mehreren Vorgängern im vorigen Zeitschritt resultieren kann.

Es gibt zahlreiche Erweiterungen (siehe Abschn. 2.3) des sRBN - Modells. Ich bezeichne sie allgemein als boolesche Zufallsnetzwerke oder *RBN*. Für einen Netzwerkzustand

$$\Sigma_t = \{\sigma_1, \sigma_2, \dots, \sigma_N\}_t \quad (2.1)$$

im Aktualisierungsschritt t bezeichne ich mit $\sigma_{i,t} \equiv \sigma_i(t)$ den Zustand des i -ten Knotens, $i = 1, \dots, N$, der die booleschen Werte 0 oder 1 annehmen kann. Bei einer festen Anzahl der Eingänge K je Knoten werden die Modelle in der Literatur auch *NK* - Modelle genannt, obwohl diese Bezeichnung eher in Verbindung mit einer *NK* - Fitnesslandschaft¹ verwendet wurde.

Die booleschen Zufallsnetzwerke haben sich erwiesen als:

- Modelle mit einem nicht trivialen reichhaltigen dynamischen Verhalten, das einer gewissen Universalität unterliegt. Wegen prinzipieller Hindernisse bei einer Simulation des dynamischen Verhaltens am Computer (siehe Diskussion der numerischen Ergebnisse in den folgenden Kapiteln) konnten viele neuere Ergebnisse (siehe Abschn. 2.4) erst seit 2003 mit Hilfe trickreicher analytischer Methoden gewonnen werden. Die vorliegende Arbeit hat hierzu einen Beitrag geleistet.

¹Bei einer *NK* - Fitnesslandschaft hängt der Fitnessbeitrag eines Knotens vom Zustand seiner K "Nachbarn" ab.

2. Überblick: Boolesche Zufallsnetzwerke (RBN)

- Ein nützliches generisches Modell für viele komplexe Systeme (siehe Abschn. 2.2). In den letzten Jahren konnte man mit Hilfe boolescher Netzwerke echte genetische regulatorische Netzwerke quantitativ simulieren. In diesem Zusammenhang ist die Untersuchung bestimmter Netzwerkkomponenten bzw. Module interessant (siehe Abschn. 2.2.5), die auch in der vorliegenden Arbeit betrieben wird.
- Modelle, an denen man topologische und funktionelle Strukturänderungen, also Netzwerkwachstum, studieren kann um unter anderem die biologische Evolution zu modellieren. Die Strukturänderungen werden in dieser Arbeit nicht betrachtet.

Man kann bei der Untersuchung der RBN - Modelle historisch grob zwei Perioden unterscheiden. Die zahlreichen Studien bis 2002 wurden in [Alda 03b] zusammengefasst. In vielen dieser Studien spielten Computersimulationen eine gewichtige Rolle, was leider zu einigen Trugschlüssen führte. Das spätere tiefere Verständnis der Modelle zusammen mit der Akkumulation der Daten über echte dynamische Netzwerke, insbesondere in der Genetik und in der Soziologie, führte zum derzeit andauernden Schub der Forschungsaktivitäten, worauf in den folgenden Abschnitten ohne Anspruch auf Vollständigkeit eingegangen wird.

2.2. Anwendungsgebiet

2.2.1. Überblick

Boolesche Zufallsnetzwerke wurden auf verschiedensten Gebieten als generische dynamische Modelle von vielen wechselwirkenden Systemelementen (Knoten) oder als ein praktisches umfassendes Konzept benutzt. Einige Beispiele sind die sozialen und finanziellen Netzwerke, die Verkehrsnetze, die neuronalen Netze, die genetischen und metabolischen Netzwerke [Alda 03b, Galp 04, Whit 86].

Boolesche Zufallsnetzwerke sind in Verbindung mit der Beschreibung der genetischen regulatorischen Netzwerke zu viel studierten Modellen geworden. Ich habe sie in diesem Zusammenhang untersucht und betone hier deshalb ihre Anwendungen in der Genetik. In Kürze und stark vereinfacht, die regulierenden Gene werden in einer Zelle transkribiert und anschließend in Proteine translatiert. Die resultierenden Genprodukte können die Produktionsrate, auch *Expression* genannt, von mehreren abhängigen Genen beeinflussen und werden auch Transkriptionsfaktoren genannt. Das Verständnis dessen, welche Gene wann exprimiert werden und was die darunter liegenden Mechanismen sind, ist der Schlüssel zur Entwicklung der Eingreifmöglichkeiten, die die Herstellung von synthetischen genetischen regulatorischen Netzwerken [Gard 00] und neue Behandlungswege für schwere Krankheiten [Huan 99] einschließen.

Boolesche Netzwerke wurden benutzt als Modelle für die genetische Regulation [Kauf 69, Kauf 74, Kauf 93, Thie 99, Thom 73, Thom 78], das Immunsys-

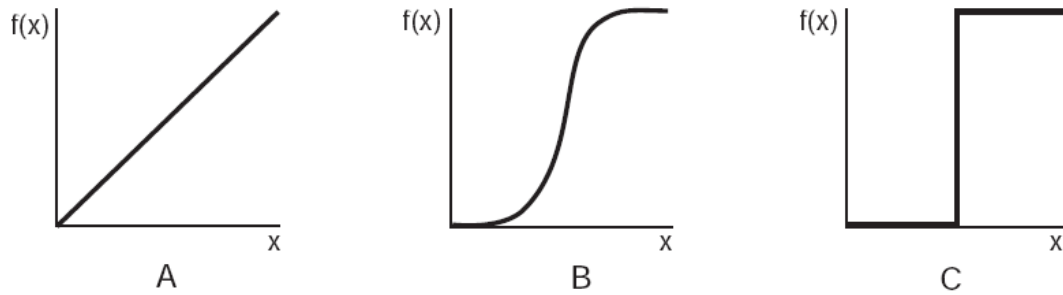


Abb. 2.1.: Sigmoidale Antwort des Systems (B) und zwei dazugehörige Näherungen, mit einer linearen Funktion (A) und mit einer Stufenfunktion (C). Die Funktion $f(x)$ kann die Produktion (Expression) eines Gens als Funktion der Expression eines aktivierenden Transkriptionsfaktors darstellen.

tem [Kauf 85, Kauf 99, Mura 96, Thie 95b], die Morphogenese [Mend 98, Sanc 97], die Metastabilität und die Homöostase [Thie 95a, Thom 91, Thom 90, Thom 95]. Boolesche Netzwerke stellen geeignete Modelle dar zur Beschreibung von Phänomenen wie globales komplexes Verhalten, Selbstorganisation, Stabilität, Redundanz, Hysterese, Periodizität, kooperatives Verhalten in biologischen Systemen [Somo 96, Szal 98, Huan 99].

2.2.2. Modellierung der genetischen Regulation

Was macht die von S. Kauffman eingeführten RBN zu einem so nützlichen Modell der genetischen Regulation? Auf der einen Seite werden sie im Vergleich zu Modellen mit kontinuierlichen Variablen durch relativ wenige Parameter beschrieben, sie können wegen der Diskretheit im Raum und in der Zeit leicht simuliert werden und erlauben einfache Vergleiche mit experimentellen Daten. Noch wichtiger, auf der anderen Seite stellen sie eine sehr gute Näherung für viele kontinuierliche zeitlich und räumlich heterogene (genetische) Systeme mit komplexem Verhalten von interagierenden Systemelementen, bei denen die nicht linearen Wechselwirkungen eine Analyse erschweren. Betrachten wir als Beispiel ein Systemelement, dessen Regulation von einem anderen Element durch eine monotone nicht lineare Funktion beschrieben werden kann, siehe Abb. 2.1. Wir interessieren uns für qualitative Änderungen des Verhaltens des ersten Elements. Näherungsweise kann man dann einen Schwellenwert für die Regulation einführen und zwei Fälle unterscheiden: die Regulation hat oder hat nicht stattgefunden. In der Genregulation findet man meistens die Situation vor, in der die Konzentration (Expression) von einem Gen als Funktion der Expression eines Transkriptionsfaktors durch eine Funktion der “sigmoidalen” [Kauf 69, Glas 73, Thom 90] Form dargestellt werden kann, siehe Abb. 2.1. In der einfachsten Näherung wird diese Regulation durch eine Stufenfunktion mit zwei “booleschen” Funktionswerten beschrieben, die mechanistischen Details

2. Überblick: Boolesche Zufallsnetzwerke (RBN)

der Regulation werden vernachlässigt. Es stellt sich heraus, dass diese Näherung die Anzahl, die Position und viele Eigenschaften der (biologisch interessanten) stationären (homöostatischen) Zustände richtig erfassen kann.

Für biologische Netzwerke geht die Idee der booleschen Beschreibung auf Jacob und Monod [Jaco 61] zurück und wurde in den 1960er und 1970er Jahren durch Arbeiten von S. Kauffman, R. Thomas, J. Sugita, L. Glass und anderen weiterentwickelt. Inzwischen ist sie erfolgreich sowohl auf kleinere genetische regulatorische Netzwerke [Yuh 98, Yuh 01, Matt 01] als auch auf ganze Genome [Thie 98]² angewandt worden.

Es gibt zahlreiche Erweiterungen des booleschen Konzepts. Es wurden mehrwertige (ganzahlige) Zustandsvariablen (mehrere Schwellen für verschiedenartige Verhaltensweisen von Genprodukten wie bei *lin-14* [Ruvk 89]) eingeführt. In [Thom 90, Snou 98] wurde die Untersuchung von in der Genetik allgegenwärtigen asynchronen Übergängen zwischen Knotenzuständen von korrelierten Knoten in die Modellbetrachtung integriert. Häufig können (nur) einzelne Vorgänge im System wie einer Zelle zusammen mit den zugehörigen Zeitskalen logisch modelliert werden, wobei Schwankungen der einzelnen Zeitverzögerungen bei Bedarf berücksichtigt werden können. Im booleschen Modell wird bei der so genannten asynchronen Aktualisierung pro Zeitschritt ein zufällig ausgewählter Knoten aktualisiert. Diese Prozedur ist sinnvoll, wenn wenig über Zeitskalen der einzelnen stochastischen regulatorischen Vorgänge bekannt ist. Ausführlicher werden verschiedene Aktualisierungsschemata im Abschn. 2.3.2 vorgestellt.

Eine weitere Ausweitung des booleschen Konzepts bilden Modelle mit dynamischer Evolution, auf die ich nicht eingehen werde. Anschaulich beschreiben sie gezielte Strukturänderungen der Netzwerke im Laufe der Zeit, mathematisch können sie als Familien polynomialer Abbildungen beschrieben werden [Andr 05].

Nicht boolesche Modelle

Der Vollständigkeit halber gehe ich auch auf alternative Modelle der genetischen Regulation ein. Einen Überblick über diese Modelle kann man sich verschaffen, indem alle Modelle danach klassifiziert werden, ob sie kontinuierliche Zustandsvariablen, diskrete bzw. boolesche Zustandsvariablen, stochastische Elemente beinhalten oder nicht³. Im Folgenden spreche ich über Modelle, die das Konzept der booleschen Näherung nicht (oder nur implizit) nutzen.

Zunächst sind die (einfachsten) linearen Modelle der Genregulation [Dhae 99, Chen 99], die der linearen Näherung in Abb. 2.1 entsprechen, nur begrenzt nützlich. Die meisten Systeme beinhalten nicht lineare Regulationen.

Die Modellierung der biochemischen Details erfolgt in der Molekularbiologie üblicherweise mit kontinuierlichen nicht linearen deterministischen “kinetischen” Differentialgleichungen [Tana 06]. Diese Beschreibung ist nur dann angemessen, wenn

²Für mehr Referenzen vgl. weiter unten im Abschn. 2.2.3.

³Grob könnte man kontinuierliche, stochastische und logische Beschreibung der genetischen Netzwerke unterscheiden.

für eine statistische Beschreibung durch mittlere Reaktionsraten genügend große Produktionsraten von relevanten Genprodukten vorliegen. Kinetische Gleichungen sind nur für kleine regulatorische Systeme praktisch nützlich bzw. handhabbar. Beim Aufstellen der kinetischen Gleichungen werden zusätzlich drastische Vereinfachungen der realen Systeme vorgenommen, beispielsweise werden bestimmte nicht transkriptionelle Vorgänge vernachlässigt. Für das Aufstellen der kinetischen Gleichungen werden detaillierte Informationen über beteiligte biochemische Prozesse benötigt. Ein berühmtes Anwendungsbeispiel aus dem Studium der embryonalen Entwicklung von *Drosophila melanogaster* [Dass 00] führte etwas überraschend zur Erkenntnis, dass die Funktion des regulatorischen Netzwerks von den biochemischen Details weitgehend unabhängig ist, vgl. im Abschn. 2.2.3 über die Anwendbarkeit des booleschen Konzepts in diesem Fall.

Um Bifurkationen im Verhalten von realen Systemen besser zu erfassen und die insbesondere in Eukaryoten wichtigen Zeitverzögerungen der Regulationen zu berücksichtigen mussten kontinuierliche Modelle mit Zeitverzögerungen [McAd 95, Chen 02] entwickelt werden. Sie können als eine Verallgemeinerung der kinetischen Gleichungen angesehen werden. Mit ihrer Hilfe kann man sowohl Umschaltvorgänge als auch kontinuierliche Änderungen im System unter Berücksichtigung der verschiedenen Zeitskalen sowohl generisch als auch anwendungsorientiert studieren.

Insbesondere bei geringen Produktionsraten von beteiligten Genprodukten müssen bekanntlich (siehe [Swai 04, Kamp 92]) Fluktuationen berücksichtigt werden, die Beschreibung muss dann notwendig durch probabilistische Modelle bzw. Gleichungen mit stochastischen Termen erfolgen. Molekulare Fluktuationen können durch Zufallsvariablen mit stochastischen Differentialgleichungen (Langevin - Gleichungen) modelliert werden [Rao 02]. Sie eignen sich besser zur Simulation, ihnen entsprechende Gleichungen für die Wahrscheinlichkeitsdichten (Fokker - Planck - Gleichungen) sind günstiger für analytische Untersuchungen [Gard 90]. Eine Alternative zur kontinuierlichen probabilistischen Beschreibung bildet die diskrete probabilistische Beschreibung durch Mastergleichungen [Rao 02], allerdings begrenzt die sehr große Anzahl möglicher zu beschreibender Zustandsübergänge ihre Anwendbarkeit.

Hybride Modelle

Viel versprechend bei der Beschreibung der genetischen Regulation sind “hybride” Modelle, die die Vorteile der booleschen Näherung der Genregulationen nutzen und sie in einem kinetischen, stochastischen oder probabilistischen Modell integrieren.

Bei geringen Produktionsraten von Genprodukten kann man diskrete Netzwerkzustände und die Übergänge zwischen ihnen im Rahmen stochastischer Petri Netze studieren [Goss 98]. Für eine probabilistische Beschreibung der stationären Zustände⁴ von kleineren Systemen unter Berücksichtigung biochemischer Details wie den Reaktionsraten sind die durch eine Hamilton - Funktion definierten Modelle wie

⁴Die stationären Zustände sind in diesen Modellen die Gleichgewichtszustände und entsprechen grob den Fixpunkten der Dynamik in diskreten booleschen Netzwerken.

2. Überblick: Boolesche Zufallsnetzwerke (RBN)

das Shea - Ackers Modell der regulatorischen *cis* - Stelle geeignet [Lago 05]. Sie gehören zur Klasse der Spinglas - Modelle der statistischen Physik.

Probabilistische Modellierung der Dynamik genetischer Netzwerke erlaubt ein Modell, das sich der Bayes'schen Netzwerke bedient [Spir 01]. Darin sind die Knotenregulationen probabilistisch definiert durch Markov'sche Ketten. Zu den Vorteilen von diesem Modell gehört, dass es auf solide mathematische Vorarbeiten aufbaut und breit anwendbar ist, insbesondere gut geeignet für "reverse - engineering" (vgl. Abschn. 2.2.3).

Parallelen zwischen booleschen Netzwerken und asymmetrischen Spinglas - Systemen, haben, insbesondere bei der Modellierung von transkriptionellen Netzwerken mit variierender Netzwerkarchitektur im Gleichgewicht, schon vor langer Zeit das Interesse der Physiker geweckt [Derr 87]. Von dieser Analogie sowie vom Studium neuronaler Netzwerke stammt ab das Nichtgleichgewicht - Modell der (Zufalls-) Schwellennetzwerke, bekannt unter dem Namen RTN (random threshold networks) [Rohl 02]. Dieses Modell zeichnet sich durch kumulative Regulationen aus, die man durch boolesche Funktionen der regulierenden Knotenwerte beschreibt, und zwar jeweils als Signum der Summe der aktivierenden minus die Summe der hemmenden Beiträge. In [Weav 99] werden ähnlich wie bei RTN verschiedenen Regulationen desselben Systemelements verschiedene reelle Gewichte zugeordnet, allerdings wird die kumulative Regulation durch nicht lineare Gleichungen beschrieben.

Die zwei Arten der Näherung in Abb. 2.1, die lineare und die boolesche, können auch zusammen verwendet werden in Form von stückweise linearen Differentialgleichungen [Glas 73, Glas 75, Glas 78, Snou 89, Jong 04, Edwa 01], die qualitativ lineare Dynamik in der Nähe der stationären Zustände der genetischen bzw. metabolischen Netzwerke und gleichzeitig umgebungs- oder fluktuationsbedingte⁵ schnelle Relaxationen in neue stationäre Zustände beschreiben. Dieser Ansatz trägt beispielsweise der Tatsache Rechnung, dass Selbsthemmung nicht durch boolesche Variablen beschreibbar ist. Er wird insbesondere verfolgt im Zusammenhang mit der Idee, genetische Regulation als (Interaktion von) Schleifen mit positiver und negativer Rückkoppelung zu betrachten [Thom 95, Thom 01a]. In jeder Schleife sind die gerichteten Genregulationen zyklisch angeordnet. Bei der positiven Rückkoppelung wird eine Zustandsänderung eines Knotens verstärkt. Die positive Rückkoppelung beschreibt (viele) metastabile Zustände bzw. differenzierte Zelltypen. Die negative Rückkoppelung beschreibt oszillierendes Verhalten bzw. Selbsthemmung, vgl. auch Abschn. 3.1. Letztendlich sind die angesprochenen Schleifen in boolescher Näherung nichts anderes als bestimmte boolesche Netzwerkrealisierungen, im Allgemeinen ausgestattet mit der neuerdings auch im Zusammenhang mit der Rückkoppelung betrachteten [Thom 01b] asynchronen Aktualisierung der Knoten. Laut [Thom 01b] ist die Beschreibung regulatorischer Netzwerke mit Hilfe von stückweise linearen Differentialgleichungen weitgehend analog zur Beschreibung mit Hilfe von

⁵Experimentell beobachtet man beispielsweise im Laufe der embryonalen Entwicklung oder der Zellteilung charakteristische quasistationäre Kontrollpunkte (checkpoints). Man ist sich nicht einig, ob die Modellierung der Zellteilung durch Ausnutzung des Konzepts der Kontrollpunkte oder der möglichen Attraktoren geschehen soll [Steu 04].

asynchronen booleschen Netzwerken, jedoch verschieden von der Beschreibung mit Netzwerken mit synchroner Aktualisierung, vgl. auch eine Gegenüberstellung von den Modellen in [Chav 06].

Sowohl kontinuierliche Gleichungen als auch die boolesche Näherung als auch stochastische Züge vereint das Modell in [Andr 06], dass explizit bestimmte Details biochemischer Mechanismen berücksichtigt. Im Modell werden boolesche Aktivierungs- und Hemmungsvorgänge durch ein nicht lineares Gleichungssystem mit stochastischen kinetischen Koeffizienten modelliert, man reproduziert qualitativ das Verhalten von klassischen RBN.

Bei den meisten Modellen des genetischen regulatorischen Netzwerks fehlt eine explizite Berücksichtigung der Umgebungsvariablen. Eine Ausnahme bildet die Arbeit [Serr 01], die den Massen- bzw. Energiefluss durch das System berücksichtigt und das Konzept boolescher Netzwerke auf kontinuierliche Zustandsvariablen inklusive denen der chemischen Substanzen der Umgebung erweitert um Degradation von organischen Substanzen zu beschreiben. Ein anderes Modell [Braz 03] ist kontinuierlich in der Zeit und beinhaltet einen regulatorischen Kern, beschrieben durch diskrete Variablen, der kontinuierliche Umgebungsvariablen bestimmt und umgekehrt.

2.2.3. Ausgewählte Anwendungen

Als Nächstes möchte ich einige Parade - Beispiele der Benutzung des Konzepts der booleschen Netzwerke für genetische regulatorische Systeme nennen, vgl. auch die Reviews [Jong 02, Jong 06].

Der boolesche Ansatz hat sich wie bereits erwähnt sowohl für kleinere genetische regulatorische Netzwerke [Davi 02, Levi 05, Yuh 98, Yuh 01] als auch für ganze Genome wie diejenigen des Bakteriums *Escherichia coli* und der Hefezelle *Saccharomyces cerevisiae* [Thie 98, Lee 02, Tong 04] bewährt. Das erste Anwendungsbeispiel des Ansatzes war wohl die bahnbrechende Beschreibung der Verbindung zwischen einfachen (logischen) transkriptionellen regulatorischen Prozessen und dynamischen Eigenschaften des *lac*-Operons in *Escherichia coli* [Mono 61]. Ein weiteres Beispiel ist die Modellierung des Zusammenspiels der Intra- und Interzellulären Wechselwirkungen im frühen Entwicklungsstadium der Fliege *Drosophila melanogaster* in [Sanc 03]. Die Autoren erhalten eine qualitativ richtige Beschreibung der Entwicklung verschiedener Phänotypen von *Drosophila melanogaster*. Im selben Organismus impliziert die Topologie des regulatorischen "segment polarity" - Netzwerks das dynamische Verhalten unabhängig von den Details der biochemischen Reaktionen [Albe 03].

In den letzten Jahren hat man sehr viel über die Struktur der realen genetischen Netzwerke erfahren [Babu 04]. Idealiert skalenfreie Topologie mit im Mittel 2 bis 3 Regulationen pro Knoten scheint die übliche Wahl der Natur für metabolische und transkriptionelle regulatorische Netzwerke zu sein [Albe 05]. Diese Topologie

2. Überblick: Boolesche Zufallsnetzwerke (RBN)

weist auf die Existenz von regulatorischen zentralen Stellen⁶ hin. Unter jeweils allen möglichen “booleschen” Genregulationen eines Gens wird in der Natur nur ein kleiner Teil realisiert, siehe Abschn. 2.3.2. Diese Erkenntnisse sind zu einem großen Teil den neuen experimentellen Microarray - Methoden zu verdanken, die mehrere Tausend Genprodukte in einer kleinen Probe gleichzeitig nachweisen können.

Auf einem wichtigen Anwendungsgebiet, der Identifizierung der Struktur der regulatorischen Netzwerke aus Genexpressionsdaten (aus cDNA - Microarray - Experimenten und aus den Daten der bereits vielen bekannten Genomsequenzierungen), auch “reverse - engineering” genannt, wurden mit verschiedenen Ansätzen große Fortschritte erzielt [Perk 06, Akut 99, Chen 99, Weav 99, Spir 01, West 06, Thom 90, Laub 04], die ausführlicheren Reviews findet man in [Bowe 01, Jong 02, Laub 04, Braz 03]. In [Lian 98] wurde eine der ersten Lösungen des inversen Problems für boolesche Netzwerke vorgestellt, der Algorithmus erlaubt es (aus wenigen bekannten Abfolgen der Netzwerkzustände) durch Berechnungen der relativen Entropie (mutual information) eines Gens und es potentiell regulierenden Gene auf die booleschen Funktionen zu schließen. Die boolesche Näherung erlaubt es hier, große Datenmengen in einem überschaubaren Rahmen zu erfassen. Ein wichtiger Schritt beim “reverse - engineering” war zu erkennen [Akut 00], dass das Rauschen in Expressionsdaten berücksichtigt werden muss, die Autoren machen das im Rahmen ihres Modells “verrauschter” boolescher Netzwerke durch Lockerung der deterministischen booleschen Regulationen mit einer gewissen Wahrscheinlichkeit.

Booleschen Modelle sind inzwischen reif für anwendungsorientierte Studien. Beispielsweise existieren professionelle Programmpakete zur Simulation von echten genetischen regulatorischen Netzwerken [Gonz 06].

Generische boolesche Modelle leisten auch wichtige Beiträge zum Verständnis von fundamentalen Themen wie Robustheit und Evolution in biologischen Systemen. Die Evolution von Netzwerken wird in dieser Arbeit nicht betrachtet, auf die Robustheit gehe ich im Folgenden kurz ein.

2.2.4. Robustheit

Der boolesche Ansatz hat zu vielen Einsichten in die Herkunft und Gesetzmäßigkeiten der Robustheit geführt, ein Thema von fundamentaler Bedeutung nicht nur in der Systembiologie.

Biologische Organismen müssen zuversichtlich funktionieren um zu überleben, sie müssen den externen (unberechenbare Umgebung) und internen (unzuverlässige Prozesse) Störfaktoren trotzen, indem sie ihre Verhaltensmuster beibehalten oder auf eine bestimmte Art anpassen. Um die Robustheit zu verstehen, werden im Review [Kita 04] relevante Fragestellungen systematisiert und durch viele Beispiele

⁶Auch die exponentielle Verteilung der Anzahl der ein- oder häufiger ausgehenden regulatorischen Verbindungen, seltener auch andere Verteilungen, wurden gefunden, manchmal ist die Wahl der am besten passenden Klasse der Verteilungsfunktionen anhand von experimentellen Daten nicht eindeutig. Jedenfalls sind die Verteilungen andersartig, insbesondere flacher, als diejenige der Zufallsgraphen.

le belegt. Ich halte mich weitgehend an diese Systematisierung bei der folgenden Beschreibung der booleschen Modelle, die sich mit der Robustheit beschäftigen.

Das mögliche Design der regulatorischen Netzwerke wird (Funktionalität beiseite) durch die Forderung nach Robustheit entscheidend eingeschränkt. Diese Einschränkung geschieht auf mehreren Ebenen und betrifft beteiligte biochemische Prozesse genauso wie “high - level” Prozesse.

Die Robustheit auf der untersten Ebene der biochemischen Prozesse äußert sich darin, dass boolesche Modelle reale Systeme beschreiben können. Die booleschen Modelle sind nach Konstruktion für die Untersuchung der Robustheit geeignet. Zu den Mechanismen, die die boolesche Beschreibung der biochemischen Prozesse begünstigen, gehört die bereits erwähnte nach dem Prinzip eines Thermostats funktionierende negative Rückkoppelung und die Bistabilität fördernde positive Rückkoppelung. Parallelen zur Kontrolltheorie sind offensichtlich. Zu prominenten Beispielen der robusten genetischen Regulationen, die durch boolesche Variablen beschrieben werden können und von den kinetischen Details der beteiligten Reaktionen weitgehend unabhängig ablaufen, gehört die Regulation der Segment - Polaritäts - Gene in *Drosophila melanogaster* [Dass 00], die Chemotaxis in *Escherichia coli* [Alon 99], das neurogene - Gene - Netzwerk in *Drosophila melanogaster* [Meir 02].

Auch auf der logischen Ebene sind die genetischen regulatorischen Netzwerke robust. Die Dynamik läuft auf wenige Fixpunkte oder kurze Attraktoren. Selbst die experimentell beobachteten Trajektorien im Zustandsraum sind häufig robust, eine Abweichung führt nach einer kurzen Zeit auf einen Netzwerkzustand dieser Trajektorie. Ausschlaggebend ist die topologische Zusammenschaltung der Knoten. Unwichtig sind dagegen die Details der Aktualisierung, Störungen etc. Als Beispiele kann man das Modelle des Segment - Polaritäts - Netzwerks in *Drosophila* [Albe 03] und den Zellzyklus in der Hefe *Saccharomyces cerevisiae* [Li 04, Brau 06] nennen. Die stabile Trajektorie des Hefe - Zellzyklus ist in Abb. 2.2 neben dem verwendeten booleschen Netzwerk dargestellt.

2.2.5. Module, Motive und Komponenten

Die im Zellzyklus der Hefe *Saccharomyces cerevisiae* in [Tong 04] genauer untersuchten 132 Gene beeinflussten im Mittel 34 der insgesamt ~ 1000 Gene des genetischen regulatorischen Netzwerks. Dieses Beispiel zeigt, dass die realen genetischen regulatorischen Systeme stark vernetzt und sehr komplex sind. Detaillierte Beschreibungen der einzelnen Genregulationen führen daher kaum zu einem tieferen Verständnis des Systemverhaltens.

Zum Glück kann man in vielen Systemen funktionale oder lokale Module von zusammenhängenden Genen identifizieren, die weitgehend unabhängig vom Rest des regulatorischen Systems sind. Für die Beschreibung dieser Module können boolesche Modelle geeignet sein. Die Module werden dann als bestimmte Netzwerkkomponenten modelliert. Die Komponenten haben nach Definition keine Verbindungen

2. Überblick: Boolesche Zufallsnetzwerke (RBN)

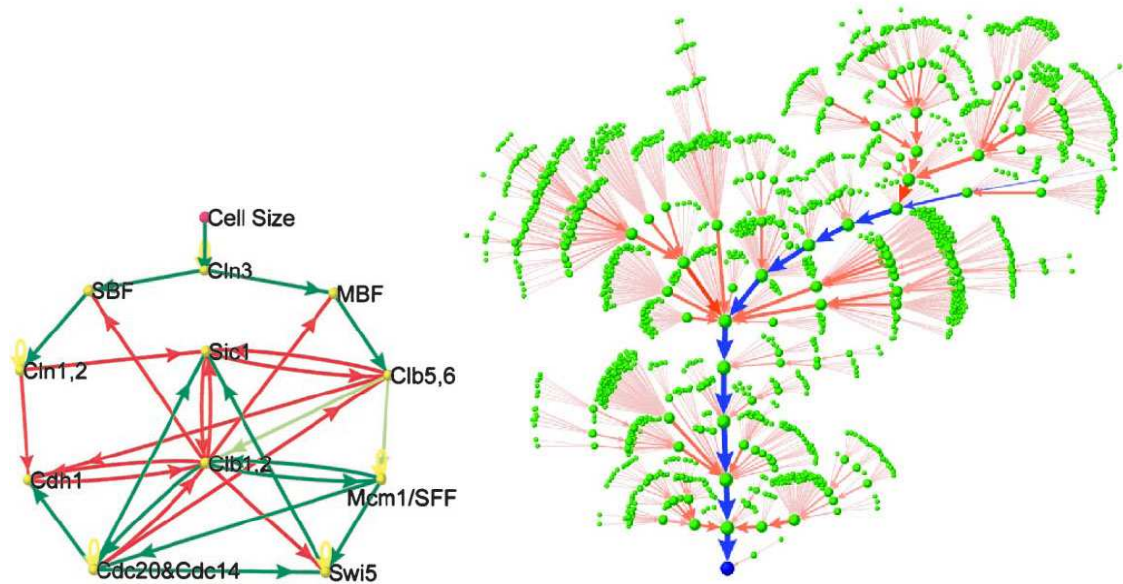


Abb. 2.2.: Im linken Bild ist das boolesche Modellnetzwerk des Hefezellzyklus (Knospung) abgebildet. Die Anzahl der jeweils vorhandenen (Knotenzustand 1) aktivierenden (grüne Pfeile) Genprodukte minus die Anzahl der vorhandenen hemmenden (rote Pfeile) Genprodukte entscheidet über den Zustand des von ihnen regulierten Knotens im nächsten Zeitschritt. Wenn diese Zahl kleiner (größer) Null ist, wird der Knoten auf 0 (1) aktualisiert, sonst behält er seinen Wert bei oder wird 0, falls er abgebaut wird (gelbe Pfeile). Im rechten Bild repräsentieren die grünen Knoten die 1764 Proteinzustände, die unter der Dynamik in den blau gefärbten Fixpunkt (*G1* - Phase) fließen. Blaue Pfeile stellen die Zellzyklus - Zustandssequenz dar. Die Abbildungen sind aus [Li 04] entnommen.

zu außen stehenden Knoten und sind innen zusammenhängend. Ein Beispiel wurde in Abb. 2.2 dargestellt.

Thomas, Thieffry und ihre Kollegen benutzten das Wort Modul in Verbindung mit periodischen Folgen von Wechselwirkungen in regulatorischen Netzwerken, siehe [Thie 99]. Diese Folgen bezeichnen die Autoren als Schleifen mit Rückkoppelung. Jeder Knoten in einer solchen Schleife beeinflusst indirekt aktivierend oder hemmend sich selbst. Man spricht von Schleifen mit positiver oder negativer Rückkoppelung. In [Snou 98] wurde gezeigt, dass man Schleifen mit positiver Rückkoppelung für Multistationarität (Zelldifferenzierung im biologischen Kontext) und Schleifen mit negativer Rückkoppelung für oszillierendes Verhalten (Oszillationen der Expression der Gene) braucht. In [Thie 99] wurde sogar angenommen, dass alle “funktionsfähigen” Schleifen mit Rückkoppelung Multistationarität oder oszillierendes Verhalten zeigen. Diese Einschränkung für das Netzwerkdesign lässt sehr große (nicht funktionsfähige) Schleifen nicht zu und impliziert eine modulare Struktur der Netzwerke.

Allgemeiner kann man von Bauelementen in genetischen regulatorischen Systemen sprechen. Schleifen mit positiver Rückkoppelung können transienten Signalen Nachhaltigkeit verleihen und als Speichermedium fungieren. Schleifen mit negativer

Rückkoppelung unterdrücken das Rauschen in der Umgebung. So genannte Feedforward - Schleifen, die Eingangssignale auf mehreren Wegen weiter propagieren lassen, können die Empfindlichkeit der Genregulationen erhöhen. Experimentell kennt man viele Module mit diesen Eigenschaften, die negative Selbstregulation kommt zum Beispiel besonders häufig vor [Hira 02].

Neben autonomen Modulen werden auch so genannte Motive als Bauelemente der Netzwerke angesehen. Der Begriff Motiv stammt aus der Sequenzanalyse in der Genetik und wurde in [Shen 02] für die genetischen Netzwerke verallgemeinert. Motive sind topologische Verbindungsmuster, die in verschiedenen Teilen des Netzwerks und dann häufiger als in Referenznetzwerken, meistens Zufallsnetzwerken, vorkommen. Über Algorithmen zur Identifizierung von Motiven kann man sich in [Newm 06] informieren. Biologisch äußern sich Motive in korrelierten Genexpressionen für entsprechende Gene. Umgekehrt geben diese Korrelationen Auskunft über die Motive im Netzwerk [Dhae 00].

Für die Motive fragt man, welche von ihnen wie häufig in realen Netzwerken auftreten. Da genetische regulatorische Netzwerke in verrauschter Umgebung sowie bei unzuverlässigen internen Mechanismen funktionieren, folgert man, dass die häufigsten Motive zur robusten Dynamik beitragen. In [Shen 02] wurden drei signifikante Motive im Genom des Bakteriums *Escherichia coli* gefunden. Die Autoren von [Lee 02] identifizierten sechs signifikante Motive im Genom der Hefezelle *Saccharomyces cerevisiae*. In [Milo 04b] wurden die Häufigkeiten der einfachsten Motive aus drei und vier Knoten in verschiedensten Netzwerken wie dem Internet und sozialen Netzen oder in genetischen regulatorischen Netzwerken bestimmt. Man fand verblüffende Ähnlichkeiten in der lokalen topologischen Struktur innerhalb verschiedener Netzwerkklassen. Im Rahmen eines booleschen Modells der in [Milo 04b] untersuchten Motive wurde gezeigt, dass ihre Häufigkeiten tatsächlich mit ihrer Fähigkeit zusammenhängen, bei Anwesenheit von stochastischen Zeitverzögerungen der Eingangssignale reproduzierbare Dynamik zu zeigen. Es hat sich außerdem herausgestellt, dass die in [Milo 04b] gefundenen Verteilungen der Motive in einem einfachen Modell zu reproduzieren schwierig ist [Milo 04a].

2.3. Modelltypen

2.3.1. Netzwerktopologie

Ich betrachte ein Netzwerk mit N Knoten mit Knotenindizes $i = 1 \dots N$. Der i -te Knoten habe k_i Eingänge und q_i Ausgänge. Eine Wahl der Verbindungen zwischen diesen Knoten definiert die Netzwerktopologie. Ein Knoten kann einen anderen Knoten oder sich selbst nur einmal als Eingang wählen. Damit das Netzwerk verknüpft ist, muss die mittlere Anzahl der Eingänge der mittleren Anzahl der Ausgänge gleich sein:

$$K \equiv \langle k \rangle = \langle q \rangle \equiv Q . \quad (2.2)$$

2. Überblick: Boolesche Zufallsnetzwerke (RBN)

Im Kontext der genetischen regulatorischen Netzwerke sind Gene mit von einigen wenigen bis von der Größenordnung 10^2 regulierenden Transkriptionsfaktoren bzw. regulierten anderen Genen bekannt [Alda 03a]. Die meisten bekannten Genregulationen erfolgen mit weniger als 5 Transkriptionsfaktoren.

In der Literatur wurden verschiedene Netzwerktopologien untersucht. Man unterscheidet die Modelle zum einen danach, wie die Nachbarschaftsbeziehungen der Knoten definiert sind.

- Die Knoten können auf einem d -dimensionalen Gitter angeordnet sein, so dass jeder Knoten seine Ein- und Ausgänge unter den $2d$ Nachbarn auswählen kann. Die Randbedingungen kann man periodisch wählen.
- Die Knoten können ihre Eingänge unter allen verfügbaren Knoten auswählen. In diesem Fall spricht man vom Kauffman'schen Modell. Ich beschränke mich in den folgenden Abschnitten auf dieses Modell. Dabei werden auch Selbstverbindungen (gerichtete Schleifen der Länge 1) zugelassen.
- Sonderfälle sind auch denkbar. Wenn mit einer gewissen Wahrscheinlichkeit p die Knoten auf einem d -dimensionalen Gitter andere Knoten als die nächsten Nachbarn als Eingang wählen, erhält man kleine - Welt - Netzwerke. In letzter Zeit wurde die kleine - Welt - Architektur in metabolischen Netzwerken gefunden [Jeon 00].

Des Weiteren ist die *Knotengrad - Verteilung* der Anzahl der Eingänge und Ausgänge $P(k, q)$ ein wichtiger Unterscheidungsmerkmal. Folgende Fälle wurden untersucht.

- Die Anzahl der Eingänge (Ausgänge) wird für jeden Knoten gleich gewählt: $k_i = K$ ($q_i = Q$).
- Meistens interessiert man sich für unabhängige Ein- und Ausgangeverteilung: $P(k, q) = P_>(k)P_<(q)$. Legt man eine der beiden Faktor-Funktionen fest und verbindet das Netzwerk ansonsten zufällig, so ergibt sich eine Poisson - Verteilung für den anderen Faktor, da jeder Knoten bei Hinzufügen der Verbindungen vom gegebenen Knoten mit derselben Wahrscheinlichkeit gewählt wird. In sRBN ist die Anzahl der Eingänge konstant und die Anzahl der Ausgänge Poisson - verteilt.
- Zunehmend im Zusammenhang mit der Beschreibung der echten Netzwerke relevante Verteilung ist die Potenzgesetz - Verteilung $P(x) \sim x^{-\gamma}$. In diesem Fall erhält man so genannte skalenfreie Netzwerke, für gewöhnlich ist $P_>(k) = P(k)$ und $P_<$ folgt einer Poisson - Verteilung. In endlichen Netzwerken muss man den maximalen Verknüpfungsgrad x_{\max} (cutoff) angeben.
- Unter den anderen Verteilungsfunktionen findet vor allem die Exponentialfunktion Anwendung.

Unterschiedliche Nachbarschaftsbeziehungen können zu qualitativ verschiedenen dynamischen Eigenschaften führen. In Netzwerken auf Gittern werden beispielsweise die Verbindungen zwischen Nachbarn bevorzugt. Das führt zu einer grundsätzlichen Bevorzugung kleiner Schleifenlängen der gerichteten topologischen Schleifen von Knoten (loops) im Vergleich zum Kauffman - Modell. Es ist dann plausibel, dass die Dynamik in den beiden Modellen unterschiedlich ist [Alda 03b]. Andererseits können verschiedene Knotengradverteilungen ähnliche oder sogar gleiche dynamische Eigenschaften implizieren. Die Ergebnisse für eine Erweiterung des sRBN - Modells aus dieser Arbeit wurden vor kurzem in [Miha 06] sogar auf allgemeine $P(k, q)$ ausgeweitet. Der Fall der skalenfreien Topologie wurde in [Deok 06, Cohe 02] untersucht⁷.

Es gibt auch andere als beschriebene Möglichkeiten für die Wahl der Netzwerktopologie. Insbesondere für neuronale Netzwerke und für Modelle von Nahrungsnetzen sind Netzwerkarchitekturen mit *gewichteten Verbindungen* interessant. Diese Gewichte können beispielsweise ein quantitatives Maß für Biomasse - Fluss zwischen zwei trophischen Spezies angeben. Weitere Möglichkeiten können ins Auge gefasst werden, nachdem für die Topologie der genetischen regulatorischen Netzwerke bekannt ist, dass sie orts-, zeit- und umgebungsabhängig sein kann [Lusc 04].

Da das dynamische Verhalten der booleschen Netzwerkensembles im Allgemeinen von der Topologie und vom gewählten Aktualisierungsschema abhängt, werden als Nächstes verschiedene Aktualisierungsschemata diskutiert.

2.3.2. Aktualisierung

Bei gegebener Netzwerktopologie wird das dynamische Verhalten von RBN durch die Angabe der Knotenfunktionen (update rules) und der zeitlichen Abfolge der Knotenaktualisierungen festgelegt.

Knotenfunktionen

Welche Knotenfunktionen mit welcher Wahrscheinlichkeit gewählt werden, hängt davon ab, was und wie modelliert wird. In dieser Arbeit habe ich hauptsächlich eine Modellierung der genetischen regulatorischen Netzwerke im Sinn. Aus diesem Grund beschränke ich mich auf die Beschreibung von bestimmten relevanten Klassen von Knotenfunktionen.

Die Anzahl der möglichen booleschen Knotenfunktionen eines Knotens mit k Eingängen ist 2^{2^k} , wobei 2^k die Anzahl der möglichen Eingänge - Zustände ist. Die Funktionen werden üblicherweise als Spalten der entsprechenden Wahrheitstafeln dargestellt, in denen für jede Kombination der Eingänge - Zustände der Ausgang angegeben wird. Für sRBN sind die 16 möglichen Funktionen in der Tabelle 2.1 aufgeführt. Eine Darstellung mit Hilfe von logischen binären Verknüpfungen ist

⁷Ich habe erfahren, dass die Quelle [Deok 06], die noch in keiner Zeitschrift veröffentlicht wurde, eine falsche Aussage enthält, die voraussichtlich bald berichtigt sein wird.

2. Überblick: Boolesche Zufallsnetzwerke (RBN)

insbesondere bei vielen Eingängen nützlich. Häufig werden dann die disjunktive (DNF) und die konjunktive (CNF) Normalform⁸ verwendet.

Wenn Genregulation von einzelnen Genen durch andere Gene als logische Schalter modelliert werden kann, dann sind die booleschen Funktionen der meisten Schalter so genannte kanalisierende Funktionen⁹. Für *kanalisierende boolesche Funktionen* fixiert zumindest ein Wert eines Eingangs den Ausgang. Die Werte der anderen Eingänge können nur dann wichtig sein, wenn dieser Eingang nicht auf seinem *kanalisierenden Wert* ist. Die Schreibweise “kanalysierend” (canalyzing) geht auf Kauffman zurück [Kauf 69]. Kanalisierende Funktionen werden auch manchmal als *forcing* (aufzwingend) bezeichnet [Stau 87]. Bei $K > 2$ hat die Anzahl der kanalisierenden Funktionen als Funktion des Anteils p der 1en in ihrer Wahrheitstafel zwei Maxima, bei einem Wert $1/2 < p_m < 1$ und symmetrisch bei $1 - p_m$. Die Anzahl und andere mathematische Eigenschaften der kanalisierenden Funktionen wurden in [Just 04] bestimmt.

Falls eine kanalisierende Funktion gänzlich unabhängig vom Wert eines Eingangs ist, kann man die entsprechende Verbindung zu dem regulierten Knoten trennen, vgl. die Klasse \mathcal{C}_1 in der Tabelle 2.1. Der Knoten hat dann effektiv weniger Eingänge. So kann ich für sRBN später von Knoten mit effektiv einem oder keinem (Klasse \mathcal{F} in der Tabelle 2.1) Eingang sprechen.

Die Funktionen, die auf alle Änderungen der Werte der Eingänge reagieren, können intuitiv zu keinem biologisch stabilen Verhalten führen und sind in realen genetischen regulatorischen Netzwerken eher eine Ausnahme. Solche Funktionen werden *reversibel* genannt (vgl. Tabelle 2.1), weil in Netzwerken mit solchen Funktionen zwei verschiedene Netzwerkzustände nach Definition nicht zum selben Folgezustand führen können, sodass eine eindeutige Umkehrung der Dynamik möglich ist. Etwas überraschend unterscheiden sich kanalisierende Funktionen nicht wesentlich von reversiblen Funktionen, siehe Abschn. 4.2 im Teil II.

Von den 140 in [Harr 02] studierten Genregulationen mit 3, 4 und 5 Eingängen entsprechen 133 den so genannten *hierarchisch kanalisierenden Funktionen* (hierarchically canalyzing functions, HCF). Diese sind eine Untermenge der kanalisieren-

⁸Ein Beispiel einer Funktion der booleschen Variablen x, y und z in CNF ist $f = (x + \bar{y}) \cdot (x + y + z)$, also eine Konjunktion von Disjunktionen, wobei $+$ für logisches ODER, \cdot für logisches UND und \bar{x} für logische Negation von x steht. Die entsprechende DNF $f = x + z \cdot \bar{y}$ ist gleichzeitig auch minimal, hat also die kleinste mögliche Anzahl der erwähnten Verknüpfungen.

⁹Beispielsweise sind 77 der 86 in der Datenbank TRANSCompel[®] der die Transkription in Eukaryoten betreffenden regulatorischen Elemente kanalisierend [Gref 06]. In [Gref 06] werden relevante biochemische Mechanismen der Genregulation in einem Modell zusammengefasst, die regulierenden Funktionen ergeben sich aus der Dynamik, in 1000 simulierten Netzwerken mit 1000 Genen hatten 91,9% der Gene kanalisierende regulatorische Funktionen. In der Arbeit von S. Harris et al. [Harr 02] hat man 139 von 140 genetischen Regulationen mit 3, 4 und 5 Eingängen als kanalisierend identifiziert. Die Autoren warnen, dass das Studium der Gene, die durch kanalisierende Eingänge reguliert werden, sich möglicherweise einfacher gestaltet oder aus anderen Gründen nicht repräsentativ sein kann. Auch vor Fehlinterpretationen der experimentellen Daten sei man nicht gefeit. Die Autoren fanden eine Tendenz zur großen Anzahl der kanalisierenden Eingänge in den Daten, mehr dazu weiter unten im Abschnitt.

den Funktionen und wurden in [Szal 98] eingeführt. Sie wurden auch in [Kauf 03] studiert, unter dem Namen *nested canalizing functions*. In jener Arbeit wurde vermutet, dass die booleschen Funktionen des regulatorischen Netzwerks vom Hefepilz *saccharomyces cerevisiae* zu HCFs gehören. Eine HCF mit k Eingängen mit Werten i_l ($l = 1, \dots, k$) wird definiert durch beliebige boolesche Werte O_{default} , I_l und O_l ($l = 1, \dots, k$), wenn der Wert des Ausgangs o beträgt $o = f_1(i_1, i_2, \dots, i_k)$, wobei $f_{k+1} = O_{\text{default}}$ und für $l = 1, \dots, k$ gilt:

$$f_l(i_l, \dots, i_k) = \begin{cases} O_l & \text{falls } i_l = I_l \\ f_{l+1}(i_{l+1}, \dots, i_k) & \text{sonst .} \end{cases}$$

Damit kein Eingang effektiv weggelassen werden kann, muss gelten $O_{\text{default}} \neq O_k$. Mit $I_1 = \dots = I_{k-1} = 0$ erhält man so genannte Kettenfunktionen (chain functions) [Gat 03]. In [Harr 02], 107 Funktionen sind Kettenfunktionen. Neulich fand man heraus [Niko 06], dass 128 der 133 HCFs in [Harr 02] den zwei einfachsten dort eingeführten Klassen von HCF angehören und für boolesche Eingänge x_1, x_2 , etc. bis auf Negierungen entweder als $x_1 \cdot x_2 \cdot x_3 \dots$ oder als $(x_1 + x_2) \cdot x_3 \dots$ geschrieben werden können.

Eine Komposition von kanalisierenden Funktionen ist im Allgemeinen nicht mehr kanalisierend. Im Teil II spielt die Tatsache eine große Rolle, dass man aus kanalisierenden Funktionen eine reversible bilden kann. Unter Komposition abgeschlossene Klassen boolescher Funktionen hat amerikanischer Mathematiker Emil Post bereits 1921 studiert [Post 21]. Die stabile Dynamik der biologischen Netzwerke kann sehr allgemein im Rahmen bestimmter Post - Klassen (A^2 und a^2) studiert werden [Shmu 03]¹⁰.

Die in [Gref 06] eingeführten *unaten* booleschen Funktionen sind eng mit der Post - Klasse der monotonen Funktionen verwandt und beschreiben 82 der 86 in der Datenbank TRANSCompel[®] zusammengestellten genetischen regulatorischen Elemente (77 sind kanalisierend) [Gref 06]. Eine boolesche Funktion ist unate, falls in jedem additiven Term ihrer minimalen DNF jede Eingangsvariable entweder selbst oder in der negierten Form vorkommt, jedoch nicht beides.¹¹

Eine Funktion aus der Menge der kanalisierenden Funktionen oder der HCFs, aus den Post - Klassen A^2 und a^2 , aus der Menge der unaten Funktionen wird jeweils wahrscheinlich einen Anteil $p \neq 1/2$ der 1en in ihrer Wahrheitstafel haben (bias), was biologisch adäquat ist.

¹⁰Die Klasse A^μ (a^μ) mit $\mu \geq 2$ beinhaltet Funktionen, die für beliebige μ oder weniger Eingangszustände, die zum Ausgang 1 (0) führen, stets einen gemeinsamen Eingang mit dem Wert 1 (0) besitzen. Ein Zusammenhang mit kanalisierenden Funktionen: Die logische Summe von $m \leq \mu$ Termen in DNF einer Funktion aus A^μ ist kanalisierend. Die Funktion $x_1 \cdot x_2 + x_2 \cdot x_3 + x_3 \cdot x_1$ ist eine Funktion aus A^2 .

¹¹Biologisch bedeutet das, dass die Eingangsvariable entweder einem aktivierenden oder einem hemmenden Transkriptionsfaktor entspricht.

Verteilung der Knotenfunktionen, Symmetrie

Nachdem die Menge der booleschen Funktionen festgelegt ist, muss man auch ihre Wahrscheinlichkeits - Verteilung angeben. Im einfachsten Fall werden alle booleschen Funktionen von K Argumenten mit der gleichen Wahrscheinlichkeit gewählt. Bei Beschränkung auf kanalisierende Funktionen werden entsprechend alle kanalisierenden Funktionen mit der gleichen Wahrscheinlichkeit gewählt. In sRBN sind die 16 booleschen Funktionen von zwei Argumenten gleich wahrscheinlich.

Man kann verschiedenen Klassen von auftretenden Funktionen eigene Wahrscheinlichkeiten zuweisen. In dieser Arbeit werden verallgemeinerte sRBN - Netzwerke betrachtet mit den Anteilen β , α , γ und Δ ($\alpha + \beta + \gamma + \Delta = 1$) der Funktionen aus den jeweiligen Klassen \mathcal{F} , \mathcal{C}_1 , \mathcal{C}_2 und \mathcal{R} , siehe Tabelle 2.1.

Im Falle der Gleichverteilung oder auch für verallgemeinerte sRBN besitzt das entsprechende Netzwerkensemble eine wichtige Symmetrie, nämlich bezüglich der Vertauschung der booleschen Werte 0 und 1. Es ist plausibel, dass diese Symmetriedynamisch stabil ist. Im Netzwerkensemble wird weder der Wert 0 noch der Wert 1 dynamisch bevorzugt. Dies ermöglicht eine Anwendung der statistischen Methoden, häufig reicht schon eine Betrachtung der Mittelwerte ohne Berücksichtigung der Fluktuationen aus, vgl. Abschn. 4.1.2. Allgemein vereinfachen vorhandene Symmetrien die Betrachtung.

Im Falle asymmetrischer Wahrscheinlichkeitsverteilungen müssen meistens andere stabilisierende Maßnahmen ergriffen werden (wie eine Veränderung der Konnektivität) um eine nicht triviale Dynamik zu erhalten.

Bei einer viel studierten Möglichkeit Knotenfunktionen zu wählen wird für jede Funktion die gleiche Wahrscheinlichkeit p des Ausganges 1 angenommen. Dadurch wird bei einer festen Anzahl k der Eingänge die Wahrscheinlichkeit für das Auftreten einer booleschen Funktion mit l Einsen in der Wahrheitstafel zu $p^l(1-p)^{2^k-l}$ und der mittlere Anteil der Einsen in der Wahrheitstafel zu p . Der Anteil p wird auch *magnetische Ausrichtung* (magnetization bias) genannt [Alda 03b]¹².

Eine andere asymmetrische Verteilung der Funktionen entstand im Kontext der neuronalen Netzwerke bzw. Perzeptrone, wo die Neuronen aktiv werden (feuern), wenn die Summe der aktivierenden und hemmenden Eingangssignale eine gewisse Schwelle h überschritten hat. Falls $\sigma_i \in \{+1, -1\}$ für aktive (inaktive) Knoten steht und die Matrixelemente $c_{ij} \in \{0, \pm 1\}$ für die abwesenden, “aktivierenden” und “hemmenden” Verbindungen der Knoten des gerichteten Graphen aus N Knoten stehen, dann werden die Knotenfunktionen wie folgt definiert: $f_i = \text{sgn}(\sum_{j=1}^N c_{ij}\sigma_j + h)$, siehe [Born 00] für den Fall $h = 0^+$. Bei einer Einschränkung auf genau $K = 2$ Eingänge sind die Funktionen im entsprechenden Netzwerkensemble die vier kanalisierenden Funktionen, die mehr Einsen als Nullen in ihrer Wahrheitstafel haben. Auch Ensembles, in denen nur ein Teil der booleschen Funktionen kumulativen Funktionen mit Schwellenwerten entsprechen wurden studiert

¹²In Anlehnung an die Analogie zwischen den Knoten und einem System wechselwirkender Spin - $1/2$ - Teilchen im ausrichtenden Magnetfeld.

Ein	\mathcal{F}		\mathcal{C}_1				\mathcal{C}_2								\mathcal{R}	
00	1	0	0	1	0	1	1	0	0	0	0	1	1	1	1	0
01	1	0	0	1	1	0	0	1	0	0	1	0	1	1	0	1
10	1	0	1	0	0	1	0	0	1	0	1	1	0	1	0	1
11	1	0	1	0	1	0	0	0	0	1	1	1	1	0	1	0

Tab. 2.1.: Die 16 booleschen Funktionen für Knoten mit zwei Eingängen. Die 4 verschiedenen Eingänge - Zustände sind in der ersten Spalte aufgelistet, die folgenden Spalten repräsentieren die 16 Funktionen. Es werden 4 Klassen von Funktionen unterschieden (vgl. Haupttext): *gefrorene* (\mathcal{F}), *kanalisierende* (\mathcal{C}_1 und \mathcal{C}_2) und *reversible* (\mathcal{R}) Funktionen.

[Volk 98]. Des Weiteren kann man allgemeiner die Koeffizienten c_{ij} als reelle Verbindungs - Gewichte zwischen -1 und 1 auffassen [Alda 03b].

Insbesondere beim Studium der Netzwerke mit gewichteten Verbindungen — sie stellen eine sehr allgemeine Klasse von Netzwerken dar — ist es sinnvoll, die Netzwerkensembles nach weiteren Kriterien wie Clusterkoeffizient (clustering coefficient) [Kahn 06], Transinformation (mutual information), mittlere Artikulation (medium articulation), gemeinsame Entropie (joint entropy) etc. zu unterscheiden [Wilh 01, Spor 02].

Bei einer biologisch begründeten Wahl der Knotenfunktionen geht man von einer Klasse wie HCF aus und wählt diejenige Verteilung der Funktionen aus, die experimentell zugängliche Größen wie den Aktivierungsgrad (die mittlere Anzahl der aktiven Gene) reproduziert [Kauf 03].

Abfolge der Aktualisierungen

Sei eine Netzwerkrealisierung (die Netzwerktopologie plus die Knotenfunktionen f_i) aus dem Netzwerkensemble gegeben. Eine Angabe der Regeln \mathcal{R} , nach denen sie ihren Zustand in einem diskreten Zeitschritt ändert,

$$\mathcal{R}(\{f_i\}, t, \Delta t) : \Sigma_t \rightarrow \Sigma_{t+\Delta t}, \quad \Delta t > 0, \quad (2.3)$$

vervollständigt das Design und definiert ein dynamisches System. Das Aktualisierungsschema in diskreten Systemen bestimmt entscheidend die dynamischen Eigenschaften.

Ich gehe im Folgenden von einem gegebenen (Anfangs-) Zustand einer Netzwerkrealisierung $\{\sigma_1, \sigma_2, \dots, \sigma_N\}_0$ aus. Im einfachsten Aktualisierungsschema werden alle Knoten gleichzeitig gemäß ihren Knotenfunktionen aktualisiert und die Werte der Knotenfunktionen definieren den neuen dynamischen Zustand des Netzwerks. Man spricht von der *synchronen Aktualisierung* (synchronous or parallel update). Diese Aktualisierung ist deterministisch, der Nachfolger - Zustand des Netzwerks ist eindeutig bestimmt. Die entsprechenden Netzwerke werden nach [Gers 04] auch als klassische RBN oder CRBN bezeichnet.

2. Überblick: Boolesche Zufallsnetzwerke (RBN)

Immer noch deterministisch aber nicht mehr synchron ist die Möglichkeit, jedem Knoten eine Verzögerungszeit zuzuweisen, nach der er jeweils aktualisiert wird. Damit werden verschiedene Zeitskalen berücksichtigt [Thom 73, Thom 78]. Die Wichtigkeit der asynchronen Aktualisierung für genetische regulatorische Netzwerke wird in [Thom 90] betont. In [Gers 04] wird der so genannte Netzwerkkontext definiert: zwei Parameter $p_i > q_i$ pro Knoten i . Ein Knoten wird dann aktualisiert, wenn $t \bmod p_i = q_i$ ergibt. Das entsprechende Netzwerkensemble wird als deterministische asynchrone RBNs bezeichnet (DARBN). Die Zeitverzögerungen beinhalten grundsätzlich neue Information für die Definition des Modells und können mathematisch mit booleschen Differenzen - Gleichungen (BDE) beschrieben werden [Dee 84]. DARBN sind nützlich für reale (charakterisiert durch refraktäre Phasen) Systeme [Kauf 99].

Die DARBN wurden im Rahmen des BDE - Formalismus auch auf den Fall kontinuierlicher Zeit (CTBN, continuous - time Boolean networks) erweitert [Okte 03]. Variierende reelle Zeitverzögerungen wurden z.B. in [Klem 05b] betrachtet.

Nach [Glas 75] geht die synchrone Aktualisierung mit einer künstlichen Einführung einer Zentraltaktsteuerung einher, die in kontinuierlichen heterogenen Systemen kaum vorkommt. Zusätzlich sind die Details einer asynchronen Aktualisierung häufig unbekannt. Daher ist eine nicht deterministische Aktualisierung sinnvoll. Mit ARBN (asynchrone RBN) wird asynchrone nicht deterministische Aktualisierung modelliert: in jedem Zeitschritt wird ein Knoten zufällig gewählt und aktualisiert [Harv 97, Grei 05]. In diesem Fall ist das Konzept der periodischen Attraktoren nicht angemessen und man greift auf *schwache Attraktoren* (loose attractors) zurück [Harv 97], die aus Netzwerkzuständen bestehen, die im Limes sehr großer Zeiten durch die Dynamik erreichbar sind. Auch im nicht deterministischen Fall können deterministische Phasenraum - Trajektorien reproduziert werden [Di P 01].

Eine interessante Realisierung der Asynchronität wurde in [Klem 05a] vorgestellt. Dort stört man ein CRBN infinitesimal und untersucht die Antwort des Systems. Somit kann man periodische Attraktoren identifizieren, die stabil gegenüber fluktuierenden Verzögerungen sind. Die Existenz von zwei im CRBN gleichzeitig ihren Wert ändernden Knoten führt bereits meistens zu einer Instabilität.

Eine weitere mögliche Definition der Asynchronität besteht darin sicherzustellen, dass alle N diskrete Zeitschritte jeder der N Knoten einmal aktualisiert wird, wobei die Reihenfolge zufällig gewählt wird oder bei Einführung einer Zeitskalentrennung für zwei nacheinander zu aktualisierende Knotengruppen zufällig gewählt wird [Chav 05]. Mit den beschriebenen Typen der Asynchronität ist es einfach möglich, das Verhalten der booleschen Netzwerke auch zwischen den zwei Grenzfällen, CRBN und ARBN, einheitlich zu studieren.

Die asynchronen Aktualisierungsschemata wie ARBN, DARBN, CTBN werden nach [Alda 03b] auch *serielle Aktualisierung* genannt, siehe [Simo 01] für ein Beispiel der seriellen Transkriptionsregulation im Hefe - Zellzyklus. Ich finde, man könnte diese Art der Aktualisierung als Aktualisierung mit Gedächtnis bezeichnen.

Nach [Gers 04] werden weitere Aktualisierungsschemata unterschieden. In verallgemeinerten asynchronen Zufallsnetzwerken (GARBN, generalized asynchronous)

werden in jedem Zeitschritt beliebig viele Knoten zufällig gewählt und aktualisiert. Die CRBN und die ARBN sind als Grenzfälle in diesem Modell enthalten, siehe [Mata 05] für ein Anwendungsbeispiel. Schließlich gibt es noch deterministische verallgemeinerte asynchrone Zufallsnetzwerke, DGARBN, die analog zu DARBN sind, mit dem Unterschied, dass alle Knoten, die $t \bmod p_i = q_i$ erfüllen, gleichzeitig aktualisiert werden.

Die Asynchronität und der Nichtdeterminismus bilden einen sehr breiten Kontext und werden benutzt um Effekte wie Zeitskalentrennung, Synchronisierung und Desynchronisierung, Stabilisierung und dergleichen zu studieren.

Es gibt einen weiteren wichtigen Aspekt. Bislang bin ich von einer konstanten Netzwerkrealisierung ausgegangen. Abgesehen von der Möglichkeit, die Netzwerkrealisierung für einen bestimmten Zweck zu evolvieren, gibt es folgende praktische Möglichkeiten, in jedem Aktualisierungsschritt die Netzwerkrealisierung zu ändern.

- Man kann trivial die einmal gewählte Netzwerkrealisierung für alle Zeiten beibehalten. Das Modell wird dann als “quenched” bezeichnet. Im “quenched” - Modell mit N booleschen Knoten hat der Zustandsraum die Größe 2^N .
- Falls in jedem Zeitschritt die Netzwerkrealisierung zufällig neu bestimmt wird (wobei der Netzwerkzustand beibehalten und dann aktualisiert wird), dann spricht man vom “annealed”¹³ - Modell oder von der Näherung “annealed approximation”. In dieser Näherung lassen sich dynamische Eigenschaften des Netzwerkensembles analytisch einfacher untersuchen [Derr 86a].
- Einige Aufmerksamkeit haben auch Verallgemeinerungen des “annealed”-Modells bekommen [Bast 96].
- In der biologisch motivierten Arbeit [Shmu 02] wurden probabilistische boolesche Netzwerke (PBN) eingeführt, in denen jedem Knoten mehrere gewichtete Funktionen zugeordnet werden, die bei Knotenaktualisierung nach bestimmten Kriterien je nach Gewicht ausgewählt werden.

In dieser Arbeit beschränke ich mich im Teil II auf deterministische synchrone “quenched” Aktualisierung in sRBN. Die Ergebnisse im Teil I sind vom Aktualisierungsschema unabhängig.

¹³Diese Begriffe kommen aus der Metallurgie. “Quenched” bezieht sich auf eine schnelle Abkühlung, “annealed” bezieht sich auf ein abwechselndes Erhitzen und Abkühlen bis das Material weniger oder keine Defekte mehr hat. Im generischen globalen Optimierungsverfahren “simulated annealing” (SA) wird in jedem Zeitschritt, der durch sinkende “Temperatur” charakterisiert wird, die vorläufige Lösung durch eine neue je nach Temperatur mehr oder weniger zufällig von der aktuellen verschiedene Lösung ersetzt, um lokale Minima des Optimierungsproblems zu meiden.

2.4. Struktur und Dynamik

In diesem Abschnitt bereite ich das Studium der kritischen booleschen Zufallsnetzwerke mit $K = 2$ Eingängen pro Knoten im Hauptteil der Arbeit vor. Ich führe relevante Begriffe am Beispiel des sRBN - Modells aus dem Abschn. 2.1 ein und fasse die bisherigen Arbeiten kurz zusammen.

Für eine Netzwerkrealisierung wird ein Anfangszustand zufällig gewählt und das Netzwerk wird iterativ synchron aktualisiert gemäß (2.3). Da der Phasenraum der Netzwerkzustände eine endliche Größe 2^N hat, erreicht das Netzwerk einen periodischen Attraktor¹⁴ der Dynamik. Die Netzwerkzustände wiederholen sich mit der Periode der Länge des Attraktors. Die Gesamtheit der Anfangszustände, die auf diesen Attraktor führen, definiert den Einzugsbereich des Attraktors. Attraktoren ohne transiente Zustände (wenn alle Zustände des Einzugsbereichs zum Attraktor gehören) werden Zyklen genannt. Die Eigenschaften der Attraktoren wie ihre Anzahl und Länge werden dynamische Eigenschaften genannt. Sie sind am interessantesten am so genannten kritischen Punkt.

Kritische Netzwerke befinden sich an der Grenze zwischen der gefrorenen und der chaotischen Phase [Derr 86a, Derr 86b]. In der gefrorenen Phase (beispielsweise bei $K < 2$) propagiert eine lokale Zustandsänderung eines Knotens im Mittel zu weniger als einem Knoten pro einen Zeitschritt. Die Attraktorlängen sind damit konstant im Limes $N \rightarrow \infty$. In der chaotischen Phase (beispielsweise im Fall, dass $K > 2$ und alle denkbaren booleschen Knotenfunktionen gleich wahrscheinlich sind) wächst ein anfänglicher kleiner Unterschied (z.B. in einem Knotenwert) zwischen zwei ansonsten identischen Netzwerkzuständen im Mittel exponentiell mit der Zeit. Mit anderen Worten wächst die Anzahl der Knoten mit unterschiedlichen Knotenwerten in den zwei Netzwerkzuständen, der so genannte Hamming - Abstand, exponentiell mit der Zeit. Eine lokale Störung breitet sich im Mittel auf mehr als einen Knoten pro Zeitschritt aus [Alda 03b].

Die Knoten eines kritischen Netzwerks können nach ihrem Verhalten auf Attraktoren klassifiziert werden. Ausführlicher wird dieses Thema im Hauptteil dieser Arbeit behandelt. Es gibt Knoten, die auf allen Attraktoren immer denselben Wert annehmen. Sie sind "gefroren". Sie sind irrelevant für die dynamischen Eigenschaften. Die auf allen Attraktoren gefrorenen Knoten bilden den gefrorenen Kern des Netzwerks. Es gibt Knoten, deren Ausgänge ausschließlich irrelevante Knoten beeinflussen (oder gar keine Knoten). Obwohl ihr Zustand sich ändern kann, sind sie irrelevant, da ihr Verhalten vollständig durch andere Knoten bestimmt wird. Die übrigen Knoten sind relevant. Sie bestimmen die Anzahl und Länge der Attraktoren. Nach Definition beeinflussen sie mindestens einen anderen relevanten Knoten.

Flyvbjerg befasste sich in [Flyv 88a] als Erster mit dem gefrorenen Kern. In [Flyv 88b] führten Flyvbjerg und Kjaer in einer analytischen Arbeit das Konzept der relevanten Knoten für Netzwerke mit einem nicht gefrorenen Eingang pro Kno-

¹⁴Auch bei asynchroner Aktualisierung kann man Attraktoren definieren als diejenigen Netzwerkzustände, die auch für sehr lange Zeiten erreicht werden können, siehe [Gers 04]

ten ($K = 1$). Die Verallgemeinerung auf kritische sRBN - Netzwerke kam von Bastolla und Parisi [Bast 96, Bast 98a]. Das Konzept der relevanten Knoten hat sich als sehr wichtig für das Verständnis der Dynamik erwiesen.

Ursprünglich hatte s. Kaufman gedacht, dass die mittlere Anzahl der Attraktoren in sRBN mit der Netzwerkgröße wie \sqrt{N} mit der Netzwerkgröße skaliert. Dasselbe Potenzgesetz wurde für Netzwerke mit kanalisierenden Funktionen erwartet [Kauf 69, Kauf 93]. Dann kamen bessere Computersimulationen und beispielsweise Bilke behauptete, dass die Anzahl der Attraktoren linear mit N ansteigt [Bilk 02]. Schließlich wurde von Socolar und Kauffman in [Soco 03] numerisch argumentiert, dass die mittlere Anzahl der Attraktoren schneller als jedes Potenzgesetz ansteigt. Sie haben auch numerisch gezeigt, dass die mittlere Anzahl der nicht gefrorenen Knoten wie $N^{2/3}$ und die mittlere Anzahl der relevanten wie $N^{1/3}$ skaliert.

Die erste und immer noch eine der schönsten analytischen Bestätigungen für diese Skalierungen wurde von Samuelsson und Troein in [Samu 03] angeführt. Allerdings sind diese Skalierung in [Samu 03] gut "versteckt", vgl. [Dros 05a], hinter mathematischen Ausdrücken. In dieser Arbeit wird eine alternative Methode entwickelt um ein intuitives vollständigeres Verständnis der Modelldynamik zu bekommen.

2. Überblick: Boolesche Zufallsnetzwerke (RBN)

Teil I.

Struktur der RBN mit $K = 2$

Experimentell sind Informationen über die Struktur der genetischen regulatorischen Netzwerke viel einfacher zugänglich als detaillierte biochemische Informationen über beteiligte dynamische Prozesse. Die Informationen über die Struktur fließen in die Definition von Modellen wie sRBN ein. Diese Struktur impliziert bereits wesentliche Züge der Dynamik. Wie erklärt sich sonst, dass verschiedenste etablierte reale Netzwerke erstaunliche Ähnlichkeiten in der Struktur aufweisen [Albe 05]?

Das sRBN - Modell ist komplex. Seine Definition im Abschn. 2.1 lässt keine einfachen Rückschlüsse auf die Dynamik zu, vgl. Abschn. 2.4. Die Dynamik kann erst verstanden werden, wenn die durch die Definition implizierte nicht triviale raumzeitliche Struktur¹⁵ aufgedeckt wird. Ein Verständnis der Dynamik wird nicht in der Sprache der Knoten und ihren lokalen Verbindungen, sondern erst in der Sprache von gewissen Klassen von den Knoten und von Netzwerkkomponenten möglich. Die in den folgenden Kapiteln betrachtete Netzwerkstruktur ist im Gegensatz zur Dynamik unabhängig vom Aktualisierungsschema.

Im Kapitel 3 beschreibe ich die auch im biologischen Kontext für sich genommen interessante Struktur der einfachsten Netzwerkkomponenten und bereite ein späteres Studium dieser Komponenten im Rahmen des Netzwerkensembles im Abschn. 4.5 und im Bezug auf ihre Dynamik im Kapitel 5 vor. Dieses Studium wird wichtige Erkenntnisse für das Verständnis des RBN - Modells bringen, siehe Kapitel 6.

Im Kapitel 4 wird die Struktur des RBN - Modells, das sRBN als Spezialfall einschließt, am kritischen Punkt untersucht. Mit Hilfe bestimmter stochastischer Prozesse wird das Netzwerkensemble im thermodynamischen Limes $N \rightarrow \infty$ studiert, relevante Skalierungen und Wahrscheinlichkeits - Verteilungen werden ermittelt, insbesondere für die Klasse der relevanten Knoten, die die Dynamik bestimmen. Das Kapitel 4 ist der wichtigste Teil dieser Arbeit, weil er die Grundlage für dynamische Untersuchungen im kritischen RBN - Modell mit $K = 2$ bildet. Dieses Kapitel ist mit einer ausführlichen Einleitung versehen.

Die folgenden Kapitel basieren auf der Arbeit von mir und den Mitverfassern von [Kauf 05a, Kauf 05b, Paul 06, Kauf 06] in der Gruppe von Prof. Barbara Drossel an der TU Darmstadt.

¹⁵Unter "Struktur" verstehe ich bestimmte Klassifizierungen der Knoten und Knotenmengen (wie nicht gefrorene Knoten) im Limes $t \rightarrow \infty$, die in diesem Sinne auch dynamische Vorgänge berücksichtigen.

3. Einfache Netzwerkkomponenten

Im sRBN - Netzwerkensemble mit einer genügend großen Anzahl der Knoten kommen beliebig komplizierte Komponenten vor, was die mögliche Topologie angeht. Diese Komponenten sind voneinander unabhängige Knotenmengen und können auch dynamisch entstehen. Ein erster Schritt zum Verständnis des sRBN - Modells ist die Identifizierung und die Charakterisierung der einfachsten möglichen Komponenten. Eine separate Untersuchung der Topologie von einfachen Netzwerkkomponenten ist ein wichtiger Teil der Untersuchung der allgemeinen Netzwerkstruktur.

Die Untersuchung der einfachen Komponenten hat auch einen anwendungsorientierten Aspekt. Boolesche Netzwerke können nämlich laut Abschn. 2.2.3 und 2.2.5 angemessen sein für eine Beschreibung der Dynamik der realen genetischen regulatorischen Systeme. Spezialisierte funktionale Module können folglich als Netzwerkkomponenten modelliert werden. Relevante Fragen sind dann, welche Netzwerkkomponenten wie oft in der Natur und im Netzwerkensemble vorkommen und wie ihre Struktur stabile Dynamik unter stochastischen Fluktuationen begünstigt, vgl. Abschn. 2.2.4.

Mit der beschriebenen Motivation wird in diesem Kapitel das Ensemble $\mathcal{E}_{N,m}$ der Netzwerke der Größe N mit einem oder zwei Eingängen pro Knoten und mit einer festen Anzahl m der Knoten mit zwei eingehenden Verbindungen betrachtet. Die Netzwerke werden analog zu sRBN jeweils zufällig verbunden. Im Kapitel 4 wird klar, dass die Teilnetzwerke der relevanten Knoten¹ in sRBN im thermodynamischen Limes $N \rightarrow \infty$ mit einer Wahrscheinlichkeit $p(m)$ zum Netzwerkensemble $\mathcal{E}_{N,m}$ mit festem $m \ll N$ gehören.

Die Netzwerke mit konstanten kleinen m stellen die einfachsten möglichen Netzwerke im Hinblick auf ihre Verflochtenheit dar. Im Folgenden werden die Fälle $m = 0, 1$ und 2 diskutiert. Bei jeweils konstantem m gehe ich auf die relativen Häufigkeiten des Auftretens der möglichen Topologien ein und zeige, dass bei $m > 0$ fast alle Netzwerke nur eine Komponente mit Knoten mit zwei Eingängen beinhalten. Zu den Ergebnissen führen kombinatorische Überlegungen.

3.1. Einfache Schleifen

Ich betrachte eine Komponente einer sRBN - Netzwerkrealisierung. Da ein Knoten effektiv weniger als 2 Eingänge haben kann (wenn seine boolesche Funktion auf

¹Die so genannten relevanten Knoten bestimmen die Dynamik in kritischen booleschen Zufallsnetzwerken, siehe Kapitel 4.

3. Einfache Netzwerkkomponenten

einen Eingang nicht reagiert oder wenn der Wert eines Eingangs sich nicht ändert), haben im einfachsten Fall alle Knoten dieser Komponente (effektiv) einen Eingang. Solche Komponenten werden einfache Schleifen genannt. Sie sind charakteristisch für Netzwerke mit $K = 1$ mit Knoten mit nur einem Eingang, die in [Flyv 88b] im Detail studiert wurden, bezüglich der Dynamik vgl. Abschn. 5.1 und [Dros 05b].

Einfache Schleifen entsprechen dem Fall $\mathcal{E}_{N,0}$: $m = 0$, kein Knoten hat mehr als einen Eingang. Die Knoten können ausschließlich in Form von gerichteten Schleifen und Bäumen², die in den Schleifen wurzeln, organisiert sein. Wie im Kapitel 4 später gezeigt wird, sind die meisten Komponenten aus dynamisch aktiven Knoten im sRBN - Ensemble einfache Schleifen mit ihren Bäumen.

Das Verhalten der Schleifen bestimmt (versklavt) das Verhalten der in ihnen wurzelnden Bäume, daher können die Schleifen allein betrachtet werden. Die Knoten in den Schleifen können eine der vier Funktionen haben: die zwei gefrorenen Funktionen mit Ausgang $f(x) \equiv 0$ oder $f(x) \equiv 1$, das Kopieren $f(x) = x$ und die Negierung $f(x) = \bar{x}$ des Eingangs x . Sobald ein Knoten eine gefrorene Funktion hat, friert die ganze Komponente ein, sodass o.B.d.A. nur die zwei letzteren Funktionen betrachtet werden können. Ein Schleife mit n Knoten und $\nu > 1$ Negierungen kann bijektiv auf eine Schleife mit n Knoten und $\nu - 2$ Negierungen abgebildet werden. Dazu markiere man den Knoten, der einer der zwei Negierungen als Eingang dient, dessen Eingangsknoten und so weiter bis einschließlich zur anderen Negierung. Dann werden die zwei Negierungen durch Kopierfunktionen ersetzt und die markierten Knoten stets durch eine “invertierende Brille” betrachtet.

Ich beschränke mich also auf Schleifen mit einer Negierung oder ohne Negierungen. Die entsprechenden Fälle der Schleifen mit einer ungeraden bzw. einer geraden Anzahl von Negierungen werden als *ungerade Schleifen* bzw. *gerade Schleifen* bezeichnet. Sie entsprechen der negativen bzw. positiven Rückkoppelung aus dem Abschn. 2.2.2 in realen genetischen regulatorischen Netzwerken. Separate unabhängige Schleifen kommen im biologischen Kontext selten vor³ und werden eher wegen der einfachen Interpretation und der anschaulichen Verbindung zwischen der Topologie und der Dynamik geschätzt. Diese Verbindung wird im Abschn. 5.1 erklärt.

3.2. Komponenten mit Verzweigungen

Ich nenne Komponenten mit $m > 0$ *komplexe Komponenten* oder *Komponenten mit Verzweigungen* und beschäftige mich hier mit den möglichen Topologien im Fall $m = 1$ und $m = 2$.

Für größere m gibt es wohl kein einfaches Verfahren um verschiedene komplexe Komponenten entsprechend ihrer Topologie und folglich auch der Dynamik zu klassifizieren. Die Anzahl der zu klassifizierenden Möglichkeiten wächst überexpo-

²Wie in der Graphentheorie bilden zusammenhängende Knoten einen Baum, wenn sie keine Schleifen bilden.

³Kurze oszillierende Schleifen findet man z.B. im regulatorischen Netzwerk des Zellzyklus von *E. coli* [Milo 04a].

nentiell mit m . Bei Bedarf könnte man jedoch eine nach gewissen Kriterien ausgezeichnete Untermenge der Netzwerke klassifizieren.

Im thermodynamischen Limes ist $m \ll N$ und die durchschnittliche Anzahl der Eingänge pro Knoten geht von oben gegen Eins. Da jeder Knoten seinen einzigen bzw. seine zwei Eingangsknoten zufällig wählt, ist die Anzahl der Ausgänge in $\mathcal{E}_{N,m}$ Poisson - verteilt mit Mittelwert nahe Eins. Daher ist es unwahrscheinlich, dass ein Knoten mit zwei Eingängen auch zwei Ausgänge hat. Ich vernachlässige solche Knoten sowie Knoten mit mehr als einem Ausgang. Die Anzahl der Knoten mit zwei Ausgängen ist dann gleich der Anzahl der Knoten mit zwei Eingängen m , beide Sorten von Knoten nenne ich *Verzweigungsknoten*, im Unterschied dazu bezeichne ich mit einer *Verzweigung* die Existenz eines Knotens mit zwei Eingängen. Wenn Bäume außer Acht gelassen werden, ergibt sich eine Komponente durch Verbinden der ihr zugehörigen Verzweigungsknoten untereinander. Die zwei Sorten von Verzweigungsknoten werden entweder direkt verbunden oder mittels Ketten von Knoten mit einem Eingang und einem Ausgang.

Zwei Komponenten mit Verzweigungen werden dann als unterschiedlich angesehen, wenn sie topologisch unterschiedlich sind. Eine unterschiedliche Anzahl von Knoten mit einem Eingang und einem Ausgang auf den Verbindungen von den Knoten mit zwei Ein- oder Ausgängen wird topologisch nicht als ein Unterschied angesehen. Topologisch ähnliche Komponenten werden in einer Klasse zusammengefasst. Dynamisch besteht für jedes Aktualisierungsschema der Hauptunterschied der Komponenten aus einer Klasse in den verschiedenen Zeitverzögerungen der Signalpropagation zwischen den Verzweigungsknoten. Die Zeitverzögerungen werden durch die Definition der Aktualisierungsregeln beeinflusst. Da zumindest biologisch motivierte Komponenten vom Aktualisierungsschema weitgehend unabhängig sein müssen, ist es plausibel, dass die Dynamik für alle Komponenten einer Klasse ähnlich sein muss. Für den Fall der synchronen Aktualisierung zeige ich exemplarisch in den Abschnitten 5.2.1 und 5.2.2, dass die vorkommenden Attraktortypen und das asymptotische Verhalten der mittleren Anzahl und Länge der Attraktoren als Funktion der Komponentengröße unabhängig sind von den Längen der Ketten zwischen den Verzweigungsknoten. Folglich beschränke ich mich auf eine Betrachtung von allen möglichen Topologien, das heißt, von allen möglichen Zusammenschaltungen der Verzweigungsknoten.

Die folgenden Ergebnisse wurden in [Kauf 06, Kauf 05a] veröffentlicht.

3.2.1. Kleine Anzahl von Verzweigungen

Nach den einfachen Schleifen sind die Komponenten mit $m = 1$ zu betrachten. Die Möglichkeit, dass der Knoten mit zwei Eingängen auch zwei Ausgänge hat wird wie gesagt vernachlässigt und es bleiben zwei topologisch verschiedene Klassen von Komponenten. Sie sind schematisch in Abb. 3.1 dargestellt. In numerischen Simulationen von $\mathcal{E}_{N,1}$ oder von sRBN finde ich die zwei Schleifen mit einer Zwischenverbindung mit der relativen Wahrscheinlichkeit $2/3$ und die Schleife mit einer Querverbindung mit der relativen Wahrscheinlichkeit $1/3$. Diese Werte können ana-

3. Einfache Netzwerkkomponenten

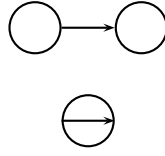


Abb. 3.1.: Schematische Darstellung der zwei topologisch verschiedenen Komponenten mit zwei Verzweigungsknoten ($m = 1$). Ein Bogen repräsentiert eine Knotenverbindung oder eine Kette von verbundenen Knoten mit einem Ein- und Ausgang. Jeder Bögen-Kreuzungspunkt steht für einen Verzweigungsknoten. Die Pfeile stellen die Eingänge der Knoten mit zwei Eingängen dar und fehlen, wenn sie für eine eindeutige Identifizierung der Komponente aus Symmetriegründen entbehrlich sind. Die erste dargestellte Komponente kommt in einem Netzwerk aus $\mathcal{E}_{N,1}$ zweimal häufiger vor als die zweite.

lytisch bestätigt werden. Für kleine m ist die Anzahl der möglichen Verbindungen der Verzweigungsknoten untereinander nämlich $(3m)!$. Man zählt dann einfach die Verbindungsweisen ab, die auf die gewünschten Komponente führen und teilt diese Anzahl durch $(3m)!$.

Analog baut man die komplexen Komponenten für $m = 2$ zusammen, siehe Abb. 3.2. Beim Abzählen ist Vorsicht geboten: Dieselbe Klasse der Komponenten kann äquivalent auf verschiedene Weisen dargestellt werden. Die erste Spalte in Abb. 3.2 ergibt bei Teilen durch $(3 \cdot 2)!$ die relativen Wahrscheinlichkeiten für das Auftreten der jeweiligen Komponente.

Die ersten drei Zeilen in Abb. 3.2 entsprechen einer Verteilung von Verzweigungsknoten auf mehrere Komponenten. Die Wahrscheinlichkeit für dieses Ereignis beträgt $72/6! = 0.1$, für $m = 1$ war sie 0. Im nächsten Abschnitt wird bewiesen, dass diese Wahrscheinlichkeit auch für allgemeine m klein ist, was später im Abschn. 4.5 benutzt wird.

3.2.2. Anzahl der komplexen Komponenten

In diesem Abschnitt zeige ich, dass die meisten Verzweigungsknoten in $\mathcal{E}_{N,m}$ sich in einer einzigen *komplexen* Komponente befinden. Die Wahrscheinlichkeit dafür, dass die Verzweigungsknoten auf mehr als eine Komponente verteilt sind verhält sich für große m wie $1/m$. Die komplexe Komponente wird maßgeblich verantwortlich für die nicht triviale Dynamik in sRBN im Vergleich zu den Netzwerken mit $K = 1$ sein, siehe Abschn. 4.5.

Für ein gegebenes m ist im Ensemble $\mathcal{E}_{N,m}$ jede Verbindung der $2m$ Verzweigungsknoten untereinander mit jedem möglichen anschließenden Einhängen der Ketten von Knoten mit einem Ein- und Ausgang in die Verbindungen gleichwahrscheinlich. Im Folgenden wird die Wahrscheinlichkeit G_m dafür berechnet, dass die $2m$ Verzweigungsknoten auf mehr als eine Komponente verteilt sind bzw. dafür, dass die $2m$ Knoten keinen zusammenhängenden gerichteten Graphen bilden. Um die

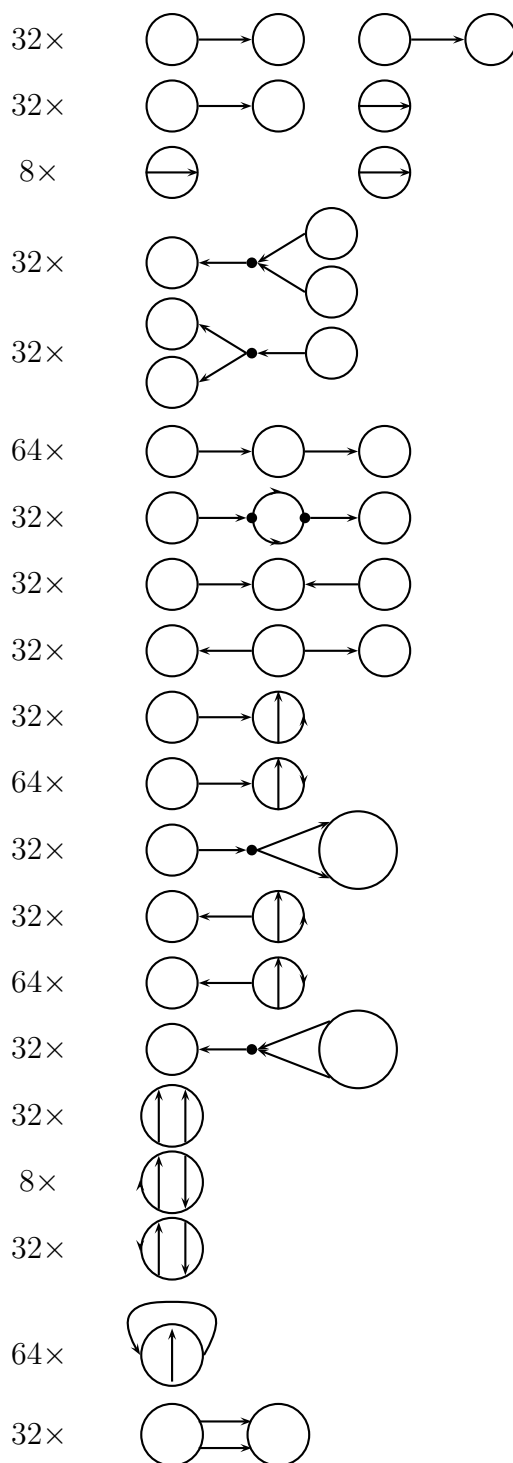


Abb. 3.2.: Schematische Darstellung der topologisch verschiedenen Komponenten für $m = 2$ Verzweigungen. Links stehen die zugehörigen Vielfachheiten. Die Bögen repräsentieren jeweils eine Knotenverbindung oder eine Kette von verbundenen Knoten. Jeder Bögen - Kreuzungspunkt steht für einen Verzweigungsknoten. • steht für genau einen Knoten. Die Pfeile stellen die Eingänge der Knoten mit zwei Eingängen dar und fehlen, wenn sie für eine eindeutige Identifizierung der Komponente entbehrlich sind.

3. Einfache Netzwerkkomponenten

entsprechende Anzahl der Verbindungen der Verzweigungsknoten untereinander, $N_m = (3m)!G_m$, zu berechnen betrachte ich irgendeinen nicht zusammenhängenden Graphen der $2m$ Knoten. Dieser Graph besteht aus mindestens zwei unabhängigen Komponenten, die eine mögliche Aufteilung der $2m$ Knoten definieren. Er trägt seine Vielfachheit zu N_m bei. Alle möglichen Aufteilungen tragen zu N_m bei. Der letzte Beitrag in der Summe kommt von der Aufteilung in m verbundene Paare von Verzweigungsknoten, einen mit zwei Eingängen und einen mit zwei Ausgängen. Jede Aufteilung wird charakterisiert durch einen Satz von ganzzahligen nicht negativen Zahlen n_i , die die Anzahl der verbundenen Gruppen aus i Verzweigungsknoten mit zwei Eingängen darstellen, wobei $m = \sum_i i n_i$ und $\sum_i n_i > 1$. Da die Anzahl der Möglichkeiten einen zusammenhängenden Graphen aus $2i$ Verzweigungsknoten zu erhalten gleich $((3i)! - N_i)$ ist, liegt eine rekursive Bestimmung von N_i nahe. Die allgemeine Formel ergibt sich unter Berücksichtigung der Vielfachheiten der Aufteilungen, die eine Doppelzählung vermeiden, zu

$$N_m = (m!)^2 \sum_{\text{Aufteilung } \{n_i|i\}} \prod_i \frac{((3i)! - N_i)^{n_i}}{n_i!(i!)^{2n_i}}, \quad (3.1)$$

mit $N_m = (3m)! G_m$.

Die Summe läuft über alle nicht trivialen Aufteilungen der m Knoten in n_i Gruppen aus i Knoten. Mit $G_1 = 0$ erhält man $G_2 = 0.1$, $G_3 = 0.1$, $G_4 \approx 0.08$, $G_5 \approx 0.07$.

Für große m kommt der größte Beitrag in der obigen Summe von der Aufteilung in ein zusammenhängendes Paar von Verzweigungsknoten und den zusammenhängenden Rest. Um das asymptotische Verhalten für große m zu untersuchen wird N_m durch diesen Beitrag angenähert. Für große m gilt demnach $G_m \approx 6m^2(3(m-1))/(3m)! \approx 2/(9m)$. Zusammenfassend befinden sich alle Verzweigungsknoten in mindestens 90% der Fälle in einer zusammenhängenden Komponente.

4. Kritische RBN mit $K = 2$

Ich beschäftige mich in diesem Kapitel mit kritischen booleschen Zufallsnetzwerken mit N Knoten mit $K = 2$ Eingängen und betrachte den thermodynamischen Grenzfall $N \rightarrow \infty$ bzw. große N . Das sRBN - Modell wird dabei als Spezialfall berücksichtigt. Der in diesem Kapitel untersuchte allgemeine Fall stellt eine Erweiterung des Standardmodells sRBN dar. *Das erweiterte Modell*, eingeführt in [Kauf 05b], wird im Folgenden verkürzt einfach *mit RBN bezeichnet*. Die Erweiterung besteht, vgl. Abschn. 2.3.2, in der Zuweisung eines relativen Gewichts für jede der 4 in Abb. 4.1 dargestellten Klassen der booleschen Funktionen.

Der Zustand der Knoten mit gefrorenen Funktionen ändert sich nicht im Laufe der Dynamik. Die Knoten mit kanalisierenden Funktionen der Klasse \mathcal{C}_1 kopieren oder invertieren einen der Eingänge, ohne auf den Zustand des anderen zu reagieren. Jeder Eingang der Knoten mit kanalisierenden Funktionen der Klasse \mathcal{C}_2 fixiert (*kanalisiert*) in einem seiner Zustände den Ausgang. Die Knoten mit reversiblen Funktionen reagieren auf jede Zustandsänderung der Eingänge. Eine Änderung eines Eingangs propagiert bei Knoten mit gefrorenen, kanalisierenden und reversiblen Funktionen jeweils mit der Wahrscheinlichkeit 0, 1/2 und 1.

In kritischen Netzwerken propagiert eine Änderung eines Knotenwerts im Mittel zu einem anderen Knoten, vgl. Abschn. 2.4. Bei den durchschnittlich 2 Ausgängen pro Knoten ist ein Netzwerk kritisch, wenn der Anteil der Knoten mit gefrorenen Funktionen gleich dem Anteil der Knoten mit reversiblen Funktionen ist. Die Anteile der Knoten mit den vier Arten von Funktionen werden bezeichnet mit

$$\begin{aligned}\alpha &= N_{c_1}/N, \\ \beta &= N_r/N = N_f/N, \\ \gamma &= N_{c_2}/N.\end{aligned}\tag{4.1}$$

Ein	\mathcal{F}		\mathcal{C}_1				\mathcal{C}_2								\mathcal{R}	
00	1	0	0	1	0	1	1	0	0	0	0	1	1	1	1	0
01	1	0	0	1	1	0	0	1	0	0	1	0	1	1	0	1
10	1	0	1	0	0	1	0	0	1	0	1	1	0	1	0	1
11	1	0	1	0	1	0	0	0	0	1	1	1	1	0	1	0

Tab. 4.1.: Die 16 booleschen Funktionen der zwei Eingänge. Die Eingänge - Zustände sind unter "Ein" aufgelistet. Es werden 4 Klassen von Funktionen unterschieden: *gefrorene* (\mathcal{F}), *kanalisierende* (\mathcal{C}_1 und \mathcal{C}_2) und *reversible* (\mathcal{R}) Funktionen.

4. Kritische RBN mit $K = 2$

Diese drei Zahlen parametrisieren das Modell.

Die Netzwerke mit $\beta = \gamma = 0$ (sie werden später als \mathcal{C}_1 Netzwerke bezeichnet) sind analog zu den kritischen $K = 1$ - Netzwerken, da die Knoten auf einen Eingang nicht reagieren, der dann “abgeschnitten” werden kann. $K = 1$ - Netzwerke unterscheiden sich von den restlichen Fällen [Flyv 88b, Dros 05b, Dros 05a] und werden hier nicht studiert. Allerdings wird sich herausstellen, dass im allgemeinen Fall das Teilnetzwerk der für die Dynamik relevanten Knoten den kritischen $K = 1$ - Netzwerken sehr ähnlich ist. Diese Analogie wird im Abschn. 4.5 ausgenutzt, dort werden einige Resultate für $K = 1$ - Netzwerke rekapituliert. (Zwecks Vergleich mit $K = 2$ kritischen Netzwerken werden \mathcal{C}_1 Netzwerke im Kapitel 6 betrachtet, unter anderem aus einem neuen Blickwinkel im Abschn. 6.1.)

Der Spezialfall der kanalisierenden Netzwerke, im Folgenden auch \mathcal{C}_2 Netzwerke genannt, $\beta = 0$, bedarf einer gesonderten Betrachtung und wird im Abschn. 4.2 diskutiert.

Im Hinblick auf das dynamische Verhalten gehören alle Netzwerke mit $\alpha < 1$ zur selben Klasse [Dros 05a]. Das dynamische Verhalten wird durch die dynamische Struktur festgelegt. Ich untersuche im Wesentlichen analytisch das Skalenverhalten der Anzahl der nicht gefrorenen (Abschn. 4.1 und 4.2) und der relevanten (Abschn. 4.3) Knoten im thermodynamischen Limes $N \rightarrow \infty$. Anschließend wird die Analogie zwischen den kritischen $K = 2$ und $K = 1$ Netzwerken betont (Abschn. 4.4). Schließlich werden die relevanten Komponenten (Abschn. 4.5) besprochen, in die die relevanten Knoten organisiert sind. Alle Ergebnisse werden durch numerische Simulationen bestätigt. Es ergibt sich ein intuitiv verständliches Bild der Struktur der kritischen RBN, das im Kapitel 6 direkt für Aussagen über die Dynamik benutzt werden kann.

Die nicht gefrorenen Knoten werden im Abschn. 4.1 behandelt. Die Definition und Analyse eines stochastischen Prozesses zur Bestimmung des gefrorenen Kerns ergibt, dass die mittlere Anzahl der nicht gefrorenen Knoten wie $N^{2/3}$ mit der Netzwerkgröße skaliert. Nur von der Größenordnung $N^{1/3}$ Knoten haben zwei nicht gefrorene Eingänge. Die zugehörigen Wahrscheinlichkeitsverteilungen werden untersucht, sie erweisen sich als bestimmte skalierte universelle Funktionen.

Im Abschn. 4.2 wird argumentiert, dass auch die kanalisierenden Netzwerke einen gefrorenen Kern der auf allen Attraktoren gefrorenen Knoten, die sich in stabilisierenden Schleifen befinden, besitzen. Die Analogie zwischen kanalisierenden und allgemeinen kritischen RBN wird anschaulich durch eine phänomenologische Theorie gezeigt.

Die relevanten Knoten werden mit Hilfe eines neuen stochastischen Prozesses im Abschn. 4.3 behandelt. Die mittlere Anzahl der relevanten Knoten divergiert bei großen N wie $N^{1/3}$. Die Wahrscheinlichkeitsverteilung zeigt Universalität im Skalenverhalten. Nur eine endliche Anzahl der relevanten Knoten besitzt zwei relevante Eingänge.

Die relevanten Knoten sind in relevante Komponenten organisiert. Die Analogie zwischen den $K = 2$ und $K = 1$ Netzwerken liefert im Abschn. 4.4 erste Aussagen über die Eigenschaften dieser Komponenten. Bis auf eine endliche Anzahl sind

im thermodynamischen Limes die relevanten Komponenten einfache Schleifen. Ihre Anzahl skaliert wie $\ln N_{rel}$ mit der Anzahl der relevanten Knoten. Die mittlere Anzahl der einfachen Schleifen der Länge $l \lesssim N_{rel}$ ist Poisson - verteilt mit dem Mittelwert $1/l$. Zusätzlich kann an dieser Stelle die im Abschn. 4.2 benutzte Aussage über die Anzahl der Knoten in stabilisierenden Schleifen in kanalisierenden \mathcal{C}_2 Netzwerken bewiesen werden.

Weiterführende Eigenschaften der relevanten Komponenten werden abschließend im Abschn. 4.5 besprochen. Das Ergebnis des Abschn. 3.2.2 kann auf das RBN - Modell direkt übertragen werden, sodass sich in mehr als 90% aller Netzwerke alle relevanten Knoten mit zwei relevanten Eingängen in einer einzigen "komplexen" Komponente befinden. Die relativen Gewichte der topologisch verschiedenen einfacheren komplexen Komponenten wurden bereits im Abschn. 3.2.1 bestimmt. Die Anzahl der relevanten Komponenten wächst logarithmisch mit der Netzwerkgröße. Je größer der Anteil der nicht gefrorenen Knoten mit zwei relevanten Eingängen, desto kleiner die Anzahl der relevanten Komponenten. Dabei wächst und wird zunehmend komplexer die größte relevante Komponente. Die Anzahl der relevanten Komponenten reduziert sich annähernd um die mittlere Größe des Logarithmus der Anzahl der relevanten Knoten mit zwei relevanten Eingängen. Die Wahrscheinlichkeitsverteilung der verschiedenen Typen der komplexen relevanten Komponenten im Netzwerkensemble wird im Limes großer Netzwerke unabhängig von der Netzwerkgröße.

4.1. Nicht gefrorene Knoten

Die Wichtigkeit einer Untersuchung des gefrorenen Kerns wurde von Flyvbjerg in [Flyv 88a] erkannt. Socolar und Kauffman haben in [Soco 03] numerisch für sRBN gezeigt, dass die mittlere Anzahl der nicht gefrorenen Knoten wie $N^{2/3}$ für große Netzwerke skaliert. Sie haben auch erste analytische Überlegungen zur entsprechenden asymptotischen Wahrscheinlichkeitsverteilung aufgestellt. Diese Wahrscheinlichkeitsverteilung wurde schließlich in dieser Arbeit (siehe auch [Kauf 05b]) bestimmt. Die ursprünglichen Ergebnisse aus [Soco 03] wurden neulich von Samuelsson und Socolar in [Samu 06] wesentlich besser verstanden, was zu einer leichten Verbesserung eines Ergebnisses aus dieser Arbeit geführt hat, wie weiter unten beschrieben.

Nachfolgend wird im Abschn. 4.1.1 ein stochastischer Prozess zur Bestimmung der auf allen Attraktoren gefrorenen Knoten in RBN beschrieben. Seine Betrachtung im Rahmen einer Mean - Field - Näherung im Abschn. 4.1.2 erweist sich als ungültig am kritischen Punkt. Durch Berücksichtigung von Fluktuationen wird das Skalenverhalten und bestimmte Grenzfälle am kritischen Punkt explizit analytisch behandelt. Zusätzlich werden die Wahrscheinlichkeitsverteilungen numerisch und implizit analytisch für die nicht gefrorenen Knoten mit einem (Abschn. 4.1.3) und mit zwei (Abschn. 4.1.4) nicht gefrorenen Eingängen bestimmt. Die Wahrscheinlichkeitsverteilungen werden durch Skalierungen von bestimmten universellen (von

den Modellparametern unabhängigen) Funktionen beschrieben. Die mittlere Anzahl der nicht gefrorenen Knoten skaliert wie $N^{2/3}$ mit der Netzwerkgröße, nur von der Größenordnung $N^{1/3}$ Knoten haben zwei nicht gefrorene Eingänge.

4.1.1. Stochastischer Prozess zur Bestimmung des gefrorenen Kerns

Flyvbjerg [Flyv 88a] war der erste, der einen dynamischen Prozess benutzt hat um ausgehend von den Knoten mit gefrorenen Funktionen iterativ den gefrorenen Kern zu bestimmen. Mit einer Mean - Field - Rechnung für diesen Prozess konnte er den kritischen Punkt identifizieren, vgl. Abschn. 2.4. Eine exakte Beschreibung des gefrorenen Kerns, die über die Mean - Field - Näherung hinausgeht, wird in dieser Arbeit durch den folgenden Prozess erreicht.

Im Netzwerkensemble, charakterisiert durch die Parameter α , β und γ und die Anzahl der Knoten N^{ini} , gehören die Knoten mit gefrorenen Funktionen trivialerweise zum gefrorenen Kern. Die Knoten des gefrorenen Kerns, deren Eingänge direkt oder indirekt durch gefrorene Knoten reguliert werden, können schrittweise identifiziert werden. In [Soco 03] werden diese Knoten “clamped” genannt.

Zur Vorbereitung der mathematischen Beschreibung werden zunächst alle Knoten gemäß ihren Funktionen auf vier Container mit den Aufschriften \mathcal{F} , \mathcal{C}_1 , \mathcal{C}_2 und \mathcal{R} verteilt. Die Knoten im Container \mathcal{C}_1 haben effektiv nur einen Eingang, den sie dynamisch entweder kopieren oder invertieren. Die Anzahl der Knoten in den Containern bezeichne ich mit N_f , N_{c_1} , N_{c_2} und N_r ; die Gesamtanzahl mit $N = N_f + N_{c_1} + N_{c_2} + N_r$. Diese Größen werden sich im Laufe des zu beschreibenden dynamischen stochastischen Prozesses verändern. Diese Änderungen werden eine “Zeitrichtung” definieren. Diese Zeit beschreibt die Zeitskala des Prozesses zur Bestimmung des gefrorenen Kerns und ist nicht mit der dynamischen Zeit für die Aktualisierungen zu verwechseln.

Die Anfangswerte der Knotenanzahl in den Containern werden mit N_f^{ini} , $N_{c_1}^{ini}$, $N_{c_2}^{ini}$ und N_r^{ini} bezeichnet, ihre Summe ergibt die Gesamtanzahl der Knoten N^{ini} . In einem Zeitschritt entfernen wir einen Knoten aus dem Container \mathcal{F} und verbinden ihn, wobei jeder Eingang der Knoten in den Containern mit derselben Wahrscheinlichkeit verbunden werden kann. Ein Knoten im Container \mathcal{C}_1 wählt den ausgezeichneten gefrorenen Knoten als Eingang mit Wahrscheinlichkeit $1/N$ und wird dann zu einem gefrorenen Knoten. Wir verschieben also jeden Knoten aus dem Container \mathcal{C}_1 mit Wahrscheinlichkeit $1/N$ in den Container \mathcal{F} . Ein Knoten im Container \mathcal{C}_2 wählt den ausgezeichneten Knoten als Eingang mit Wahrscheinlichkeit $2/N$ und wird mit Wahrscheinlichkeit $1/2$ zu einem gefrorenen Knoten (der ausgezeichnete Knoten ist mit dieser Wahrscheinlichkeit kanalisierend für den \mathcal{C}_2 - Knoten) und ansonsten zu einem \mathcal{C}_1 - Knoten. Deshalb verschieben wir jeden Knoten aus dem Container \mathcal{C}_2 in einem Zeitschritt jeweils mit Wahrscheinlichkeit $1/N$ in die Container \mathcal{F} und \mathcal{C}_1 . Schließlich wird ein Knoten im Container \mathcal{R} mit Wahrscheinlichkeit $2/N$ zu einem \mathcal{C}_1 - Knoten. Alle \mathcal{R} - Knoten werden in einem Zeitschritt mit Wahrscheinlichkeit

$2/N$ in den Container \mathcal{C}_1 verschoben. Die Gesamtanzahl der Knoten verringert sich in einem Zeitschritt um Eins, da wir den ausgezeichneten Knoten verbunden und aus seinem Container entfernt haben.

Im nächsten Zeitschritt wird ein weiterer gefrorener Knoten aus dem Container \mathcal{F} verbunden. Dadurch wird sein Einfluss auf die anderen Knoten berücksichtigt. Diese Prozedur kann wiederholt werden, bis entweder der Container \mathcal{F} leer wird oder bis alle anderen Container leer sind. Im zweiten Fall friert das gesamte Netzwerk ein und die Dynamik läuft immer auf einen Fixpunkt. Im ersten Fall bleiben in den Containern die nicht gefrorenen Knoten übrig. Anschließend können die nicht gefrorenen Knoten untereinander zu einem nicht gefrorenen Netzwerk (zufällig) verbunden werden.

Wir werden sehen, dass die meisten übrig gebliebenen Knoten im \mathcal{C}_1 Container sind und dass der relative Anteil der Knoten in den Containern \mathcal{C}_2 und \mathcal{R} im Limes $N^{ini} \rightarrow \infty$ verschwindet.

4.1.2. Mean - Field - Betrachtung

Der im letzten Abschnitt eingeführte Prozess wird zuerst mit Hilfe von deterministischen Gleichungen beschrieben, die Fluktuationen der mittleren Änderung der Knotenanzahl in den Containern werden vernachlässigt. Diese Näherung ist angemessen, solange die Anzahl der Knoten in jedem Container viel größer als Eins ist. Die mittleren Änderungen der Knotenanzahl in den Containern betragen in einem Zeitschritt

$$\begin{aligned}\Delta N_r &= -\frac{2N_r}{N}, \\ \Delta N_{c_2} &= -\frac{2N_{c_2}}{N}, \\ \Delta N_{c_1} &= -\frac{N_{c_1}}{N} + \frac{N_{c_2}}{N} + \frac{2N_r}{N}, \\ \Delta N_f &= -1 + \frac{N_{c_1}}{N} + \frac{N_{c_2}}{N}, \\ \Delta N &= -1.\end{aligned}\tag{4.2}$$

Die Anzahl der Knoten in den Containern N wird in jedem Zeitschritt um 1 kleiner und kann daher anstelle der Zeit verwendet werden. Wir nähern die Differenzengleichungen durch kontinuierliche Differentialgleichungen mit Ableitungen nach der Zeit N an und können nach N_r auflösen:

$$\frac{\Delta N_r}{\Delta N} \simeq \frac{dN_r}{dN} = \frac{2N_r}{N},$$

sodass

$$N_r = N^2 \frac{N_r^{ini}}{(N^{ini})^2}.\tag{4.3}$$

4. Kritische RBN mit $K = 2$

Analog erhält man

$$\begin{aligned} N_{c_2} &= N^2 \frac{N_{c_2}^{ini}}{(N^{ini})^2}, \\ N_f &= N \frac{N_f^{ini} - N_r^{ini}}{N^{ini}} + N^2 \frac{N_r^{ini}}{(N^{ini})^2}, \\ N_{c_1} &= N \frac{N_{c_1}^{ini} + N_{c_2}^{ini} + 2N_r^{ini}}{N^{ini}} - N^2 \frac{2N_r^{ini} + N_{c_2}^{ini}}{(N^{ini})^2}. \end{aligned} \quad (4.4)$$

Bei der Lösung sind drei Fälle zu unterscheiden.

- Bei $N_f^{ini} < N_r^{ini}$ wird $N_f = 0$ bei einem positiven $N \sim N^{ini}$ erreicht. Ein nicht verschwindender Anteil der nicht gefroren Knoten entspricht der chaotischen Phase.
- Bei $N_f^{ini} > N_r^{ini}$ werden N_r und N_{c_2} beide kleiner 1, sobald N auf den Wert größenordnungsmäßig $\sqrt{N^{ini}}$ abfällt. Für kleinere N sind nur die zwei Container \mathcal{F} und \mathcal{C}_1 nicht leer und der zweite Term in den Formeln für N_f und N_{c_1} in (4.4) kann vernachlässigt werden. Wenn N_f auf einen Wert unter 1 sinkt, verbleiben noch $N_{c_1} = \frac{N_{c_1}^{ini} + N_{c_2}^{ini} + 2N_r^{ini}}{N_f^{ini} - N_r^{ini}}$ Knoten im Container \mathcal{C}_1 . Das Netzwerk friert im Wesentlichen ein, da im Limes $N^{ini} \rightarrow \infty$ nur ein verschwindender Anteil der Knoten nicht gefroren ist. Durch verbinden der \mathcal{C}_1 - Knoten untereinander erhält man einfache Schleifen und in ihnen wurzelnde Bäume. Diese Eigenschaft der gefrorenen Phase ist auch aus [Soco 03] bekannt.
- Dem Fall der kritischen Netzwerke $N_f^{ini} = N_r^{ini} = \beta N^{ini}$ gilt mein Hauptinteresse.

Im kritischen Fall endet der stochastische Prozess laut (4.4) bei $N_f = 1 = \beta N^2 / N^{ini}$ und die Anzahl der nicht gefrorenen Knoten ergibt sich zu

$$N^{end} = \sqrt{\frac{N^{ini}}{\beta}}. \quad (4.5)$$

Sie würde also in der deterministischen Näherung mit der Wurzel aus der Netzwerkgröße skalieren. Im nächsten Abschnitt wird gezeigt, dass man bei Berücksichtigung der Fluktuationen diesen Exponenten von $1/2$ auf $2/3$ korrigieren muss. Deterministisch beträgt die Anzahl der \mathcal{C}_2 - Knoten am Ende des Prozesses am kritischen Punkt γ/β unabhängig von N^{ini} . Die Endanzahl der Knoten im Container \mathcal{R} hat wegen $N_r = N_f$ einen Wert kleiner 1. Mit Fluktuationen wird die Endanzahl der Knoten in den Containern \mathcal{C}_2 und \mathcal{R} mit $(N^{ini})^{1/3}$ skalieren.

Im kritischen Fall $N_r^{ini} = N_f^{ini}$ vereinfachen sich die Ausdrücke in (4.4). Mit den Bezeichnungen $n = N/N^{ini}$ und $n_j = N_j/N^{ini}$ für $j = r, f, c_1, c_2$ erhält man

$$\begin{aligned} n_r &= \beta n^2 = n_f, \\ n_{c_2} &= \gamma n^2, \\ n_{c_1} &= n - 2\beta n^2 - \gamma n^2. \end{aligned} \quad (4.6)$$

Folglich ist der stochastische Prozess in der deterministischen Näherung invariant gegenüber einer Multiplikation der Anfangswerte für die Anzahl der Knoten in allen Containern mit einem gleichen konstanten Faktor. Bei einem beliebig kleinen festen n befinden sich die meisten Knoten wegen $n_{c_1} = n - \mathcal{O}(n^2)$ im Container \mathcal{C}_1 . Falls man dann ein genügend großes N^{ini} wählt, nimmt n einen beliebig kleinen Wert an, während $N_f = N_r = \beta n^2 N^{ini}$ immer noch groß genug für die deterministische Beschreibung ist. Aus diesem Grund wird im Folgenden angenommen, dass für genügend große Netzwerke n und damit $N_f/N = \beta n$ kleine Größen im System werden, bevor die Fluktuationen wichtig werden. Diese Annahme vereinfacht die folgende Betrachtung.

4.1.3. Skalierung unter Berücksichtigung der Fluktuationen

Fokker - Planck - Gleichung

Wir betrachten den stochastischen Prozess aus dem Abschn. 4.1.1 zum Zeitpunkt, wenn die Fluktuationen des abfallenden Wertes N_f um den Mittelwert wichtig werden. Wegen (4.6) ist dann für genügend große N^{ini} der Wert von $n = N/N^{ini}$ beliebig klein.

Die Anzahl der Knoten im Container \mathcal{C}_1 , die den jeweils ausgezeichneten gefrorenen Knoten als Eingang wählen ist Poisson - verteilt mit dem Mittelwert gleich der Varianz gleich N_{c_1}/N . Für genügend große N^{ini} gilt $N_{c_1}/N = 1 - (2\beta + \gamma)n = 1 - \mathcal{O}(n) \approx 1$. Die Anzahl der Knoten in den Containern \mathcal{C}_2 und \mathcal{R} , die den ausgezeichneten gefrorenen Knoten als Eingang wählen ist auch Poisson - verteilt mit dem Mittelwert und der Varianz $2(N_{c_2} + N_r)/N$, siehe (4.2). Die entsprechenden Fluktuationen können vernachlässigt werden, da dieses Rauschen klein ist im Vergleich zu N_r und N_{c_2} , deren Werte am Ende des stochastischen Prozesses bei genügend großen N^{ini} viel größer als die Fluktuationen sind. Die stochastische Version der Gleichungen (4.2) lautet dann

$$\begin{aligned}\Delta N_r &= -\frac{2N_r}{N}, \\ \Delta N_{c_2} &= -\frac{2N_{c_2}}{N}, \\ \Delta N_f &= -\frac{N_r}{N} - \frac{N_f}{N} + \xi, \\ \Delta N &= -1.\end{aligned}\tag{4.7}$$

Der Mittelwert der Zufallsvariable ξ ist gleich 0, die Varianz ist gleich 1. Solange die relative Änderung der Anzahl der Knoten in den Containern n_j in einem Zeitschritt klein ist, kann man in der Gleichung für ΔN_f eine große Anzahl T Zeitschritte zu einem effektiven Zeitschritt zusammenfassen. Das Rauschen wird dann Gauß - verteilt mit dem Mittelwert 0 und der Varianz T . Man hätte denselben Prozess bei einem Zusammenfassen von T Zeitschritten eines analogen Prozesses mit einem Gauß'schen Rauschen der Varianz 1 erhalten. Folglich können wir annehmen, dass die Zufallsvariable ξ Gauß - verteilt mit Varianz 1 ist.

4. Kritische RBN mit $K = 2$

Im Vergleich zur deterministischen Näherung haben sich die Gleichungen für N_r und N_{c_2} nicht geändert, sodass immer noch gilt

$$\begin{aligned} N_r &= N^2 N_r^{ini} / (N^{ini})^2, \\ N_{c_2} &= N^2 N_{c_2}^{ini} / (N^{ini})^2. \end{aligned} \quad (4.8)$$

Einsetzen in die Gleichung für N_f liefert

$$\frac{dN_f}{dN} = \frac{N_f}{N} + \frac{\beta N}{N^{ini}} + \xi, \quad (4.9)$$

wobei die Schrittgröße $dN = -1$ und $\langle \xi^2 \rangle = 1$ ist. Wir bekommen eine Langevin - Gleichung! Im Limes $dN \rightarrow 0$ erhält man übrigens die Rauschen - Autokorrelationsfunktion $\langle \xi(N) \xi(N') \rangle = \delta(N - N')$.

Zur Herleitung der der Langevin - Gleichung entsprechenden Fokker - Planck - Gleichung sei $P(N_f, N)$ die Wahrscheinlichkeit, N_f Knoten im Container \mathcal{F} zu finden, wenn die Gesamtzahl der Knoten in allen Containern N beträgt. Diese Wahrscheinlichkeit hängt (siehe (4.9)) von N_{ini} und β ab.

Die Summe

$$\sum_{N_f=1}^{\infty} P(N_f, N) \simeq \int_0^{\infty} P(N_f, N) dN_f$$

ist die Wahrscheinlichkeit dafür, dass der stochastische Prozess noch nicht beendet ist ($N_f \neq 0$), als die Gesamtzahl der Knoten in den Containern auf den Wert N abfällt. Die absorbierende Randbedingung $P(0, N) = 0$ berücksichtigt, dass nach Erreichen von $N_f = 0$ die Systeme aus dem betrachteten Ensemble entfernt werden.

Als Nächstes wird mit $g(\Delta N_f | N_f, N)$ die Wahrscheinlichkeit bezeichnet, dass bei gegebenen N_f und N die Größe N_f sich während des nächsten Zeitschritts um $-\Delta N_f$ verringert. Sodann

$$\begin{aligned} P(N_f, N - 1) &= \int_0^{\infty} P(N_f + \Delta N_f, N) g(\Delta N_f | N_f + \Delta N_f, N) d(\Delta N_f) \\ &= \int_0^{\infty} \left[P(N_f, N) g(\Delta N_f | N_f, N) \right. \\ &\quad + \frac{\partial}{\partial N_f} (P(N_f, N) g(\Delta N_f | N_f, N)) \Delta N_f \\ &\quad + \frac{\partial^2}{2 \partial N_f^2} (P(N_f, N) g(\Delta N_f | N_f, N)) (\Delta N_f)^2 \\ &\quad \left. + \dots \right] d(\Delta N_f) \\ &= P(N_f, N) + \frac{\partial}{\partial N_f} (P(N_f, N) \langle \Delta N_f \rangle) + \\ &\quad \frac{\partial^2}{2 \partial N_f^2} (P(N_f, N) \langle (\Delta N_f)^2 \rangle) + \dots \end{aligned}$$

Der Mittelwert der Änderung $\langle \Delta N_f \rangle$ während eines Schrittes beträgt $\langle \Delta N_f \rangle = \frac{N_f}{N} + \frac{\beta N}{N^{ini}}$, die mittlere quadratische Abweichung ist $\langle (\Delta N_f)^2 \rangle \simeq 1$. Hieraus folgt die Fokker - Planck - Gleichung für den stochastischen Prozess zur Bestimmung des gefrorenen Kerns:

$$-\frac{\partial P}{\partial N} = \frac{\partial}{\partial N_f} \left(\frac{N_f}{N} + \frac{\beta N}{N^{ini}} \right) P + \frac{1}{2} \frac{\partial^2 P}{\partial N_f^2}. \quad (4.10)$$

Skalenverhalten

Mit den Skalenvariablen

$$x = \frac{N_f}{\sqrt{N}} \text{ und } y = \frac{N}{(N^{ini}/\beta)^{2/3}} \quad (4.11)$$

und der Funktion $f(x, y) = (N^{ini}/\beta)^{1/3} P(N_f, N)$ transformiert man die Fokker - Planck - Gleichung schließlich in die Form

$$y \frac{\partial f}{\partial y} + f + \left(\frac{x}{2} + y^{3/2} \right) \frac{\partial f}{\partial x} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} = 0. \quad (4.12)$$

Wir werden sehen, dass die Funktion $f(x, y)$ von den Modellparametern N^{ini} und β unabhängig ist.

Die Wahrscheinlichkeit $W(N)$ dafür, dass N nicht gefrorene Knoten in den Containern übrig bleiben, wenn $N_f = 0$ wird, ist

$$W(N) = \int_0^\infty P(N_f, N) dN_f - \int_0^\infty P(N_f, N-1) dN_f.$$

Daraus folgt

$$\begin{aligned} W(N) &\simeq \frac{\partial}{\partial N} \int_0^\infty P(N_f, N) dN_f \\ &= (N^{ini}/\beta)^{-1/3} \frac{\partial}{\partial N} \sqrt{N} \int_0^\infty f(x, y) dx \\ &= (N^{ini}/\beta)^{-2/3} \frac{\partial}{\partial y} \sqrt{y} \int_0^\infty f(x, y) dx \\ &\equiv (N^{ini}/\beta)^{-2/3} G(y) \end{aligned} \quad (4.13)$$

mit der hier eingeführten Skalenfunktion $G(y)$.

Die Funktion $W(N)$ muss normiert sein: $\int_0^\infty W(N) dN = \int_0^\infty G(y) dy = 1$. Da diese Bedingung von den Modellparametern unabhängig ist, sind es die Funktionen $G(y)$ und $f(x, y)$ auch. Das rechtfertigt die Wahl des Vorfaktors in der Definition von $f(x, y)$.

Das Skalenverhalten der mittleren Anzahl der gefrorenen Knoten folgt aus

$$\bar{N} = \int_0^\infty N W(N) dN = (N^{ini}/\beta)^{2/3} \int_0^\infty G(y) y dy. \quad (4.14)$$

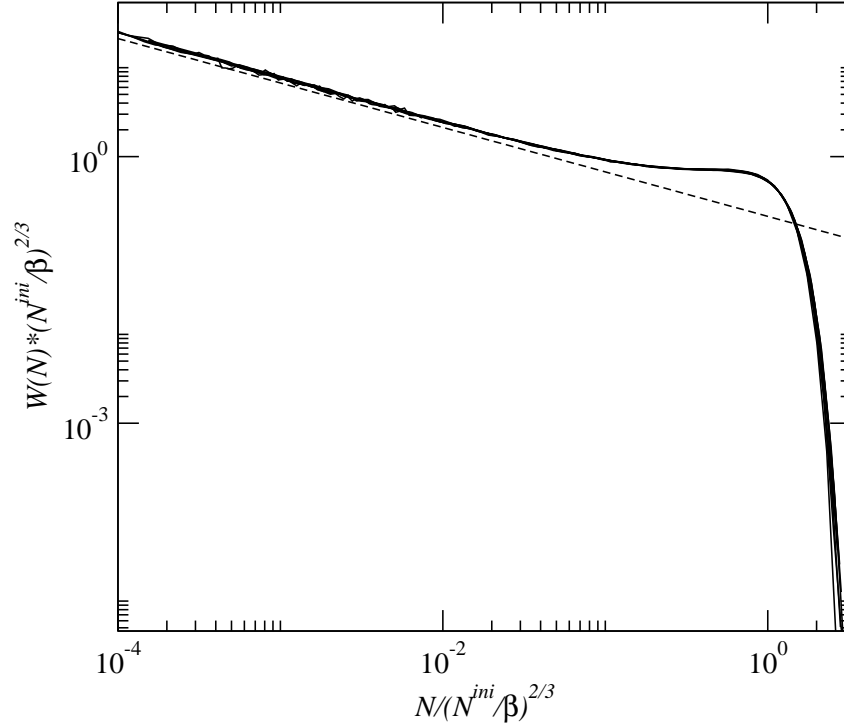


Abb. 4.1.: Dargestellt ist $W(N)(N^{ini}/\beta)^{2/3}$ als Funktion von $N/(N^{ini}/\beta)^{2/3}$ für $\beta = 0.25$ und $N^{ini} = 2^{16}, 2^{17}, 2^{18}, 2^{19}, 2^{20}, 2^{21}$. Zusätzlich sind die Kurven für $\beta = 0.125, N = 2^{16}$ und für $\beta = 0.5, N = 2^{16}$ dargestellt. Die Existenz der universellen Skalenfunktion $G(y)$ wird durch das Zusammenfallen der Kurven klar bestätigt. Die gestrichelte Linie zeigt das Potenzgesetz $1/(4\sqrt{y}) \sim 1/\sqrt{N}$.

Bemerkenswert, alle Ergebnisse in diesem Abschnitt hängen nicht von α und γ ab und die Abhängigkeit von β und N^{ini} wird durch bestimmte Skalierungen von universellen Funktionen beschrieben.

Wahrscheinlichkeitsverteilung

Ich komme zur genaueren Diskussion der Funktionen $G(y)$ und $f(x, y)$. Ein expliziter Ausdruck für die Funktion $G(y)$ konnte nicht gewonnen werden, allerdings können bestimmte Grenzfälle analytisch untersucht werden, zusätzlich können die Wahrscheinlichkeitsverteilungen mit dem eingeführten stochastischen Prozess effizient numerisch simuliert werden und führen zu weiteren analytischen Ausdrücken.

Das in Abb. 4.1 präsentierte Ergebnis für $G(y)$ wurde in numerischen Simulationen des durch (4.7) beschriebenen stochastischen Prozesses bestimmt. Ein fast perfektes Fit liefert der Ausdruck

$$G(y) \simeq 0.25e^{-y^3/2}(1 - 0.5\sqrt{y} + 3y)/\sqrt{y}. \quad (4.15)$$

Für kleine y ergeben die Simulationen das Potenzgesetz $G(y) \propto y^{-1/2}$. Diese Abhängigkeit kann analytisch bewiesen werden. Dazu wird die Fokker - Planck -

Gleichung (4.12) im Limes kleiner y gelöst. In diesem Grenzfall kann der zu $y^{3/2}$ proportionale Term vernachlässigt werden:

$$y \frac{\partial f}{\partial y} + f + \frac{x}{2} \frac{\partial f}{\partial x} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} = 0. \quad (4.16)$$

Eine allgemeine Lösung hat die Form $f(x, y) = \sum_{\nu} c_{\nu} y^{\nu} f_{\nu}(x)$ mit den Funktionen f_{ν} , die der folgenden Gleichung genügen müssen:

$$2(\nu + 1)f_{\nu} + x f'_{\nu} + f''_{\nu} = 0. \quad (4.17)$$

Die Lösung lautet:

$$e^{\frac{x^2}{2}} f_{\nu}(x) = C_1 H_{1+2\nu} \left(\frac{x}{\sqrt{2}} \right) + C_2 {}_1F_1 \left(-\nu - \frac{1}{2}; \frac{1}{2}; \frac{x^2}{2} \right),$$

C_1 und C_2 sind zwei Konstanten, H bezeichnet die Hermit - Polynome, ${}_1F_1$ ist die Standardbezeichnung für die entsprechende hypergeometrische Funktion. Da $f(x, y)$ für kleine y eine analytische Funktion von y sein muss, kann $\nu = 0, 1, 2, \dots$ gefolgert werden. Für genügend kleine y liefert nur der Term $\nu = 0$ einen nicht verschwindenden Beitrag. Aus der absorbierenden Randbedingung folgt ferner $C_2 = 0$. Für kleine y ergibt sich

$$f(x, y) = c_0 x e^{-x^2/2}. \quad (4.18)$$

Die Integration der Gleichung (4.12) über x von 0 bis Unendlich führt wegen $f(0, y) = f(\infty, y) = 0$ zu

$$\sqrt{y} \frac{\partial}{\partial y} \sqrt{y} \int_0^{\infty} f dx - \frac{1}{2} \frac{\partial f}{\partial x} \Big|_{x=0} = 0,$$

es ergibt sich ein Zusammenhang zwischen $G(y)$ und $f(x, y)$:

$$\sqrt{y} G(y) = \frac{1}{2} \frac{\partial f}{\partial x} \Big|_{x=0}. \quad (4.19)$$

Eine Kombination der numerischen Formel (4.15) und des analytischen Resultats (4.19) führt zu $c_0 = 0.5$. Einsetzen von (4.18) in (4.13) liefert die Wahrscheinlichkeitsverteilung der Anzahl der nicht gefrorenen Knoten für kleine N :

$$W(N) = \left(\frac{N^{ini}}{\beta} \right)^{-1/3} \frac{c_0}{2\sqrt{N}}. \quad (4.20)$$

Die Funktion $f(x, y)$ in (4.18) hängt nicht von y ab. Daher hängt die Funktion $P(N_f, N)$ für genügend kleine N nur vom Quotienten N_f/\sqrt{N} ab. Dieses Resultat wurde in Computersimulationen bestätigt, siehe Abb. 4.2.

Eine Lösungsmenge der Gleichung (4.12) kann implizit durch den Ansatz $f(x, y) = \sum_{\nu} y^{\nu} \tilde{f}_{\nu}(z)$ mit $z = x - y^{3/2}$ angegeben werden. Die resultierende Gleichung für \tilde{f}_{ν}

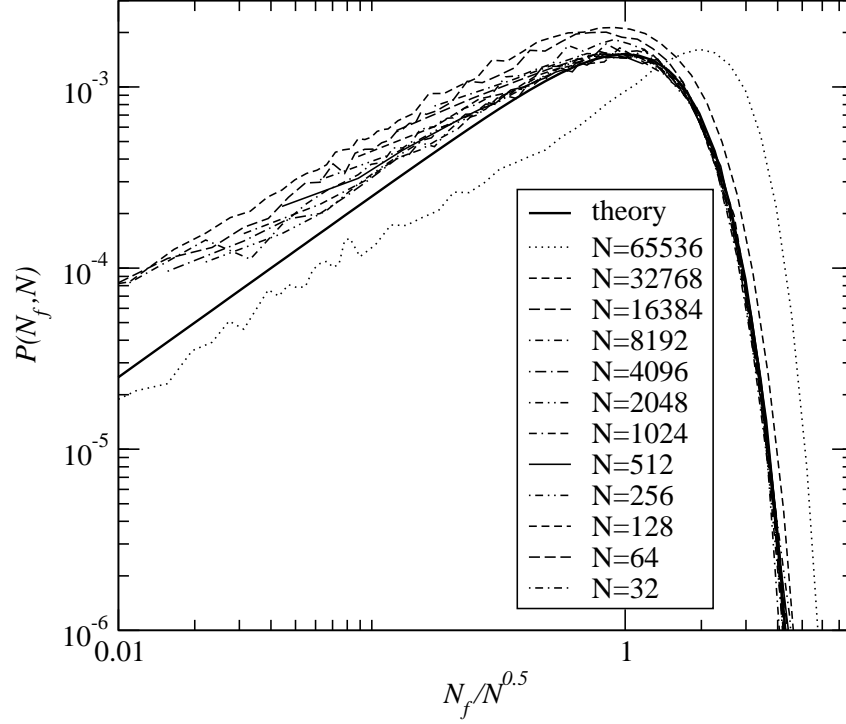


Abb. 4.2.: $P(N_f, N)$ als Funktion von N_f/\sqrt{N} für $N^{ini} = 2^{21}$ und $\beta = 1/4$ für verschiedene N . Die dicke durchgezogene Kurve ist das analytische Ergebnis (4.18), das im Grenzfall kleiner $N/(N^{ini})^{2/3}$ erreicht wird.

ist identisch mit der Gleichung (4.17) für f_ν , die für kleine y aufgestellt wurde. Allerdings kann ein analytischer Ausdruck für $G(y)$ nur dann angegeben werden, wenn die Entwicklung der Anfangsbedingung $P(N_f, N^{ini}) = \delta(N_f - \beta N^{ini})$ nach den Lösungen bestimmt werden kann.

Neben dem Grenzfall kleiner y kann auch der Grenzfall großer y für $G(y)$ analytisch untersucht werden. Durch eine anspruchsvolle Analyse eines anderen als hier stochastischen Prozesses konnte man in [Samu 06] für große y den folgenden asymptotischen Ausdruck analytisch gewinnen:

$$G(y) \approx \sqrt{\frac{2y}{\pi}} \exp(-y^3/2). \quad (4.21)$$

Des Weiteren wurde in [Samu 06] ein verbessertes numerisches Fit für $G(y)$ angegeben, das für große y besser passt:

$$G(y) \approx \sqrt{\frac{2y}{\pi}} \exp(-y^3/2) \left(1 + \frac{1}{3.248y + 4.27y^2 + 4.76y^3} \right). \quad (4.22)$$

Der maximale relative Fehler beträgt dabei 0.25% im Vergleich zu 6% bei (4.15).

4.1.4. Nicht gefrorene Knoten mit zwei relevanten Eingängen

Die Wahrscheinlichkeit $W_r(N_r)$, dass am Ende des im Abschn. 4.1.1 beschriebenen stochastischen Prozesses, wenn der Container \mathcal{F} leer wird, N_r Knoten im Container \mathcal{R} übrig bleiben, beträgt aus (4.8)

$$N_r = N^2 N_r^{ini} / (N^{ini})^2.$$

Mit der Skalenvariablen

$$s = \frac{N_r}{(N^{ini}/\beta)^{1/3}} = y^2$$

und der Skalenfunktion

$$F(s) = \frac{G(\sqrt{s})}{2\sqrt{s}} \quad (4.23)$$

folgt aus $W(N)dN = W_r(N_r)dN_r$ der Ausdruck für die Wahrscheinlichkeitsverteilung der Anzahl der nicht gefrorenen Knoten mit reversiblen Funktionen

$$W_r(N_r) = (N^{ini}/\beta)^{-1/3} F(s). \quad (4.24)$$

Die mittlere Anzahl von Knoten im Container \mathcal{R} am Ende des Prozesses skaliert wie

$$\begin{aligned} \bar{N}_r &= \int_0^\infty W_r(N_r) N_r dN_r = (N^{ini}/\beta)^{1/3} \int_0^\infty s F(s) ds \\ &= (N^{ini}/\beta)^{1/3} \int_0^\infty y^2 G(y) dy. \end{aligned} \quad (4.25)$$

Die in numerischen Simulationen bestimmte Skalenfunktion $F(s)$ ist in Abb. 4.3 dargestellt.

Die Endanzahl der Knoten im Container \mathcal{C}_2 folgt mit (4.8) aus den Ergebnissen für N_r :

$$N_{c_2} = (\gamma/\beta) N_r. \quad (4.26)$$

Die meisten nicht gefrorenen Knoten haben also einen nicht gefrorenen Eingang. Die Ergebnisse für das Skalenverhalten werden exakt im thermodynamischen Limes $N^{ini} \rightarrow \infty$. Das dynamisch interessante Teilnetzwerk aus nicht gefrorenen Knoten erhält man am Ende des stochastischen Prozesses durch Verbinden der nicht gefrorenen Eingänge. Dieses reduzierte Netzwerk wird im Abschn. 4.3 studiert.

4.2. Nicht gefrorene Knoten in kanalisierenden Netzwerken

Kanalisierende Netzwerke haben viele Eigenschaften gemeinsam mit anderen kritischen Netzwerken. Eine Analyse der Anzahl von Attraktoren in sRBN in [Samu 03] impliziert nach [Dros 05a] auch für kanalisierende Netzwerke (neben den anderen

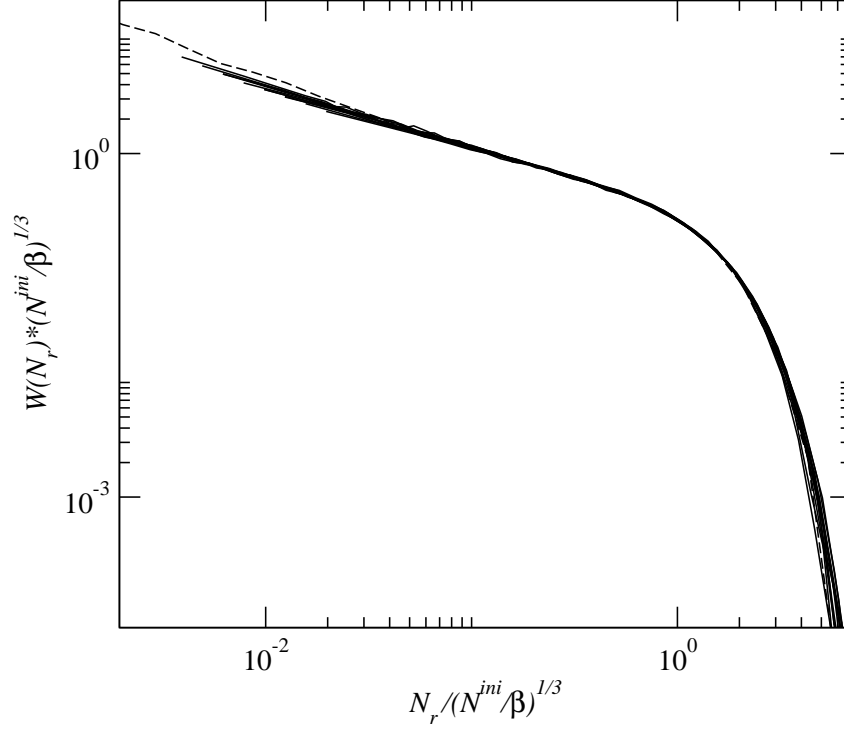


Abb. 4.3.: Dargestellt ist $W(N_r)(N^{ini}/\beta)^{1/3}$ als Funktion von $N_r/(N^{ini}/\beta)^{2/3}$ für $\beta = 0.5$ und $\beta = 0.125$ und für $N^{ini} = 2^{16}, 2^{17}, 2^{18}, 2^{19}, 2^{20}, 2^{21}$. Die 12 Kurven nähern sich der asymptotischen Form bei steigenden Werten von N . Die Existenz der universellen Skalenfunktion $F(s)$ wird durch das Zusammenfallen der Kurven bestätigt. Die gestrichelte Kurve zeigt die Daten für $F(s)$, die mit (4.23) aus den Daten für $G(y)$ aus jeweils derselben Simulation stammen. Sie zeugt von der Konsistenz unserer Analyse.

kritischen Netzwerken mit $K = 2$) die mittlere Anzahl $\sim N^{2/3}$ von nicht gefrorenen Knoten und nicht mehr als $\sim N^{1/3}$ nicht gefrorene Knoten mit zwei nicht gefrorenen Eingängen.

Allerdings lassen sich die detaillierten Ergebnisse des Abschn. 4.1 zunächst wegen Fehlens von Knoten mit gefrorenen Funktionen nicht auf die kanalisierenden Netzwerke übertragen. Außerdem ist aus den bisherigen Arbeiten nicht klar, wie gefrorene Knoten in kanalisierenden Netzwerken entstehen.

Ich argumentiere auf der Basis der in [Paul 06] veröffentlichten Ergebnisse, dass die kanalisierenden Netzwerke einen gefrorenen Kern aus allen Attraktoren gefrorenen Knoten besitzen (Abschn. 4.2.1). Eine weitgehende Analogie zwischen kanalisierenden und allgemeinen kritischen RBN wird durch eine phänomenologische Theorie im Abschn. 4.2.2 erklärt. Diese Theorie liefert unabhängig das Ergebnis für die Größe des gefrorenen Kerns. Weitere Ergebnisse zu kanalisierenden Netzwerken findet man im Kapitel 6.

4.2.1. Stabilisierende Schleifen

Obwohl es in kanalisierenden Netzwerken keine Knoten mit gefrorenen Funktionen gibt, haben diese \mathcal{C}_2 Netzwerke einen gefrorenen Kern. Ein intuitives Verständnis seiner Entstehung wird in diesem Abschnitt aufgebaut.

Da von vornherein keine Knoten für alle Anfangsbedingungen gefroren sind, müssen Gruppen von Knoten existieren, die ihre Zustände gegenseitig festlegen und auf Zustandsänderungen von Knoten außerhalb der jeweiligen Gruppe nicht reagieren. Das einfachste Beispiel einer solchen Gruppe ist eine gerichtete Schleife aus Knoten, die jeweils den Zustand ihres Nachfolgers kanalisieren (fixieren), sobald sich ihr eigener Wert auf dem Mehrheitsbit (der in der Wahrheitstafel der Knotenfunktion 3-mal vorkommt) eingependelt hat. Diese Schleifen werden einfache “stabilisierende” Schleifen genannt. Allgemeiner können Knoten in stabilisierenden Schleifen keine Eingänge von außerhalb bekommen.

Eine einfache stabilisierende Schleife bleibt allein wegen der Wahl der Knotenfunktionen für immer gefroren, wenn einmal all ihre Knoten auf ihrem Mehrheitsbit sind. Für einen beliebigen Anfangszustand kann nun die Anzahl der Knoten einer stabilisierenden Schleife, die auf ihrem Mehrheitsbit sind, konstant bleiben oder zunehmen. Sie bleibt konstant nur im unwahrscheinlichen Fall, dass die Knoten in der Schleife immer die nicht kanalisierenden Eingänge von außerhalb bekommen. Man kann daher annehmen, dass zumindest größere stabilisierende Schleifen auf allen Attraktoren des Netzwerks einfrieren.

Wenn in einem kanalisierenden Netzwerk beide Eingänge eines Knotens in stabilisierenden Schleifen sitzen oder wenn ein Eingang dieses Knotens von ihnen kanalisiert wird, friert der Knoten ein. Solche Knoten können ein Einfrieren von weiteren Knoten verursachen und so weiter. Dabei entstehen auch Knoten, die effektiv nur einen Eingang haben, weil der andere konstant ist. An dieser Stelle kann man aber eine neue Art von “verallgemeinerten stabilisierenden Schleifen” konstruieren, die Ketten von Knoten mit effektiv einem Eingang (\mathcal{C}_1 Knoten) zwischen sich gegenseitig kanalisierenden Knoten mit kanalisierenden \mathcal{C}_2 Funktionen (\mathcal{C}_2 Knoten) einschließen. Wenn eine Kette zwischen zwei \mathcal{C}_2 Knoten als Ganzes ihren Eingang invertiert, muss das invertierte Mehrheitsbit des ersten \mathcal{C}_2 Knotens den zweiten \mathcal{C}_2 Knoten kanalisieren.

Wie die einfachen stabilisierenden Schleifen sind die verallgemeinerten stabilisierenden Schleifen meistens auf allen Attraktoren gefroren. Diese Schleifen und die Knoten, dessen Einfrieren sie verursachen bilden den gefrorenen Kern von kanalisierenden Netzwerken.

4.2.2. Skalierung der Anzahl gefrorener Knoten

Ein stochastischer dynamischer Prozess kann analog zu demjenigen für RBN außer den kanalisierenden Netzwerken im Abschn. 4.1 dazu benutzt werden, die Größe des gefrorenen Kerns zu bestimmen.

4. Kritische RBN mit $K = 2$

Ich beginne mit einfachen stabilisierenden Schleifen und bestimme iterativ die Knoten, deren Einfrieren sie verursachen. Die Anzahl der Knoten, die im Laufe dieses Prozesses bereits eingefroren sind und deren Einfluss auf die anderen Knoten noch zu untersuchen ist, bezeichne ich mit f . n_1 sei die Anzahl der Knoten, die momentan genau einen im Prozess gefrorenen Eingang, der sie nicht kanalisiert, haben. Das sind die Knoten mit effektiv einem Eingang, sie werden in diesem Abschnitt \mathcal{C}_1 Knoten genannt. n_2 sei die Anzahl der Knoten (genannt \mathcal{C}_2 Knoten), die (noch) keinen gefrorenen Eingang bekommen haben. Die Anfangsbedingungen lauten: $n_1 = 0$, $f = f_0$, $n_2 = N - f_0$ und $n \equiv f + n_1 + n_2 = N$, N ist die Netzwerkgröße. Im Abschn. 4.4.1 wird übrigens bewiesen, dass für große Netzwerke die Anzahl der Knoten in einfachen stabilisierenden Schleifen $f_0 \sim N^{1/3}$ ist. Allerdings ist die Kenntnis dieses Wertes im Folgenden nicht wesentlich.

In jeweils einem "Zeitschritt" des dynamischen Prozesses wird für einen ausgezeichneten gefrorenen Knoten ausgewürfelt, wessen Eingang er wird. Er wird ein Eingang eines der n_2 Knoten mit Wahrscheinlichkeit $2/n$. Dadurch wird der \mathcal{C}_2 Knoten mit jeweils Wahrscheinlichkeit $1/2$ gefroren oder er wird zu einem \mathcal{C}_1 Knoten mit effektiv einem Eingang. Ein \mathcal{C}_1 Knoten wählt den ausgezeichneten gefrorenen Knoten als Eingang mit Wahrscheinlichkeit $1/n$ und friert dann ein. In jedem Schritt wird ein gefrorener Knoten verbunden und von der folgenden Betrachtung ausgeschlossen. Die Änderung der Anzahl von gefrorenen Knoten f in einem Zeitschritt wird mit Δf bezeichnet, die Änderung von n_1 und die Änderung von n_2 pro Zeitschritt wird mit Δn_1 und Δn_2 bezeichnet. Danach gilt:

$$\begin{aligned}\Delta f &= -1 + \xi_1 + \xi_2, \\ \Delta n_1 &= (n_2 - n_1)/n, \\ \Delta n_2 &= -2n_2/n.\end{aligned}\tag{4.27}$$

ξ_1 und ξ_2 bezeichnen Poisson - verteilte Zufallsvariablen mit den Mittelwerten n_1/n und n_2/n . Sie berücksichtigen jeweils Fluktuationen der Anzahl der vom ausgezeichneten gefrorenen Knoten in einem Zeitschritt gewählten \mathcal{C}_1 und \mathcal{C}_2 Knoten. Ihre Kombinationen in der zweiten und dritten Gleichung wurden durch entsprechende Mittelwerte ersetzt, da die Fluktuationen klein im Vergleich zu den Werten von n_1 und n_2 sind. Diese Fluktuationen werden erst bei kleinen Werten von f wichtig, wobei sich f bei einer wiederholten Auswertung der Gleichungen verringert. Die Summe $n = f + n_1 + n_2$ verringert sich genau um Eins in jedem Schritt. Der dynamische Prozess wird beendet, wenn alle Knoten gefroren sind (was unwahrscheinlich ist) oder wenn es keine unverbundenen gefrorenen Knoten mehr gibt.

Simulationen von diesem Prozess zeigen, dass die Gesamtzahl der von den einfachen stabilisierenden Schleifen eingefrorenen Knoten ungefähr mit $N^{0.8}$ skaliert¹, während von der Größenordnung N Knoten von den stabilisierenden Schleifen nicht eingefroren werden. Diese Anzahl der gefrorenen Knoten ist zu klein um den gefrorenen Kern zu erklären, da sie im Widerspruch zu direkten Simulationen von

¹Das asymptotische Verhalten wird bei diesen Simulationen erreicht.

kanalisierenden Netzwerke bzw. zu den analytischen Ergebnissen von [Dros 05a] steht. Folglich müssen wir zusätzlich die am Ende des Prozesses entstandenen verallgemeinerten stabilisierenden Schleifen aus dem Abschn. 4.2.1 berücksichtigen.

Von den n am Ende des beschriebenen Prozesses übrig gebliebenen Knoten haben n_1 Knoten Funktionen aus der Klasse \mathcal{C}_1 . n_2 Knoten haben \mathcal{C}_2 Funktionen. Nach Abschn. 4.1.3 bzw. aus Simulationen $n_1 \sim n \sim N$. Daher gibt es fast in jeder Schleife aus übrig gebliebenen Knoten einen oder mehr \mathcal{C}_1 Knoten. Die einzige dynamische Auswirkung der \mathcal{C}_1 Knoten in verallgemeinerten stabilisierenden Schleifen besteht in einer Verzögerung der Signalübertragung zwischen angrenzenden \mathcal{C}_2 Knoten. Die verallgemeinerten stabilisierenden Schleifen können daher als einfache stabilisierende Schleifen im effektiven \mathcal{C}_2 Netzwerk aus den n_2 Knoten mit \mathcal{C}_2 Funktionen angesehen werden.

Die Anzahl der Knoten in stabilisierenden Schleifen im effektiven \mathcal{C}_2 Netzwerk beträgt $\sim n_2^{1/3}$ analog der obigen Anmerkung zu f_0 . Wir bezeichnen diese Anzahl erneut mit f_0 um die Analogie zur entsprechenden Anfangsbedingung des Prozesses (4.27) zu betonen. Die f_0 Knoten werden als gefroren angesehen und die Analogie ist für die momentanen Werte n_1 , n_2 und n als Anfangsbedingungen im effektiven \mathcal{C}_2 Netzwerk perfekt, weil die notwendigen Korrekturen der Werte n_1 , n_2 und n (auf Grund der Tatsache, dass einige dieser Knoten den stabilisierenden Schleifen angehören) für große Netzwerke verschwindend klein im Vergleich zu den Werten n_1 , n_2 und n selbst sind. Schließlich können wir den Prozess (4.27) erneut starten und die Anzahl der Knoten bestimmen, deren Einfrieren durch die neuen f_0 Knoten verursacht wird.

Fortan interessieren wir uns nur für die Änderungen von n_1 und n_2 . Die Gleichungen

$$\begin{aligned}\Delta n_1 &= (n_2 - n_1)/n, \\ \Delta n_2 &= -2n_2/n\end{aligned}\tag{4.28}$$

gelten nun für die zwei beschriebenen Prozesse zusammen. Für n gilt weiterhin $\Delta n = -1$ in jedem Zeitschritt.

Wir nähern (4.28) durch entsprechende Differentialgleichungen für dn_1/dn und dn_2/dn an. Die Lösung ergibt sich zu

$$n_2 = \frac{n^2}{N},\tag{4.29}$$

$$n_1 = n - \frac{n^2}{N}.\tag{4.30}$$

Wie man schon ahnen kann, ergibt sich am Ende des zweiten Prozesses wieder ein effektives \mathcal{C}_2 Netzwerk mit Ketten aus neu erzeugten \mathcal{C}_1 Knoten, es gibt neue verallgemeinerte stabilisierende Schleifen. Es stellt sich die Frage, ob immer wieder (4.28) verwendet werden kann.

Im ersten, zweiten, und so weiter Prozess verringert sich die Anzahl der übrig gebliebenen \mathcal{C}_2 Knoten, die Anzahl der \mathcal{C}_1 Knoten steigt an. Somit nimmt das relative Gewicht der \mathcal{C}_1 Knoten im nicht gefrorenen Teilnetzwerk zu. Diese Tendenz

4. Kritische RBN mit $K = 2$

und die Iteration der Prozesse (4.28) bricht in sich zusammen, wenn die verbleibenden nicht gefrorenen Knoten nicht mehr als ein effektives \mathcal{C}_2 Netzwerk angesehen werden können. Das wird wahr, wenn das relative Gewicht n_2/n der \mathcal{C}_2 Knoten größenordnungsmäßig $1/\sqrt{n}$ wird. Eine Erläuterung folgt.

Am Ende eines der obigen Prozesse wird ein Versuch unternommen, in einem (großen) kanalisierenden Netzwerk aus dem entsprechenden Ensemble eine verallgemeinerte stabilisierende Schleife zusammenzubauen. Wenn $n_2/n \sim 1/\sqrt{n}$ gilt, dann werden mit Wahrscheinlichkeit $\sim n_1/n$, also bevorzugt, \mathcal{C}_1 Knoten ans offene Ende der wachsenden Knotenkette angeschlossen, bis ihre Länge größenordnungsmäßig \sqrt{n} beträgt. Danach ist die Wahrscheinlichkeit einen \mathcal{C}_2 Knoten anzuschließen, $\sim 1/\sqrt{n}$, von derselben Größenordnung wie die Wahrscheinlichkeit einen der Knoten aus der Kette ans Ende der Kette anzuschließen, sodass die Kette wahrscheinlich zu einer Schleife aus \mathcal{C}_1 Knoten wird. Stabilisierende Schleifen werden dagegen unwahrscheinlich.

Ich bezeichne mit N_{nf} und mit N_2 jeweils die mittlere Anzahl der nicht gefrorenen Knoten und die mittlere Anzahl der nicht gefrorenen Knoten mit zwei nicht gefrorenen Eingängen in \mathcal{C}_2 Netzwerken. $N_1 = N_{nf} - N_2$ ist dann die mittlere Anzahl der Knoten mit einem nicht gefrorenen Eingang. Die Iteration der obigen Prozesse kann nicht mehr fortgesetzt werden, wenn $N_2 \sim \sqrt{N_{nf}}$. Einsetzen von $n_2 \sim \sqrt{n}$ in (4.29) ergibt schließlich

$$\begin{aligned} N_{nf} &\sim N^{2/3}, \\ N_2 &\sim N^{1/3}. \end{aligned} \tag{4.31}$$

Das anschauliche Verständnis des gefrorenen Kerns im Abschn. 4.2.1 wird damit bestätigt. In Übereinstimmung mit [Samu 03, Dros 05a] besteht der gefrorene Kern aus allen bis auf von der Größenordnung $N^{2/3}$ Knoten, $N^{1/3}$ haben zwei nicht gefrorene Eingänge. Diese Ergebnisse für die nicht gefrorenen Knoten sind für kanalisierende und andere kritische RBN daher gleich. Im Übrigen können die Überlegungen aus diesem Abschnitt auch für gemischte RBN mit allen kanalisierenden Knotenfunktionen, sowohl aus der Klasse \mathcal{C}_1 als auch \mathcal{C}_2 , wiederholt werden, falls die Anzahl der \mathcal{C}_2 Knoten mit zwei nicht gefrorenen Eingängen größenordnungsmäßig \sqrt{N} übersteigt. Andernfalls erhält man effektiv ein \mathcal{C}_1 Netzwerk.

4.3. Relevante Knoten

Die Erkenntnis, dass die relevanten Knoten die Dynamik im thermodynamischen Limes bestimmen, war ein wichtiger Schritt zum Verständnis von RBN. Das Konzept der relevanten Knoten haben Flyvbjerg und Kjær [Flyv 88b] für $K = 1$ Netzwerke eingeführt. Auf allgemeine kritische Netzwerke wurde es von Bastolla und Parisi [Bast 98b, Bast 98a] angewendet, die die modulare Struktur von relevanten Elementen numerisch untersuchten. Socolar und Kauffman zeigten in [Soco 03], dass die mittlere Anzahl der relevanten Knoten wie $N^{1/3}$ skaliert. Dasselbe Ergeb-

nis verbirgt sich, wie von Drossel in [Dros 05a] gezeigt, hinter dem analytischen Studium der Anzahl der Attraktoren von Samuelsson und Troein in [Samu 03].

Nachfolgend wird im Abschn. 4.3.1 ein stochastischer Prozess zur Bestimmung der relevanten nicht gefrorenen Knoten beschrieben. Als Nächstes wird im Abschn. 4.3.2 gezeigt, dass die mittlere Anzahl der relevanten Knoten bei großen N mit $N^{1/3}$ skaliert. Die zugehörige Wahrscheinlichkeitsverteilung wird analytisch und numerisch bestimmt. Schließlich argumentiere ich im Abschn. 4.3.3, dass nur eine endliche Anzahl der relevanten Knoten zwei relevante Eingänge besitzt. Die entsprechende Wahrscheinlichkeitsverteilung wird analytisch bestimmt.

4.3.1. Stochastischer Prozess zur Bestimmung relevanter Knoten

Anfangen von den Knoten mit gefrorenen Funktionen beziehungsweise von den verallgemeinerten stabilisierenden Schleifen aus dem Abschn. 4.2.1 im Falle der kanalisierenden Netzwerke² kann man wie im Abschn. 4.1 beschrieben den gefrorenen Kern und die Eigenschaften der nicht gefrorenen Knoten mit Hilfe eines stochastischen Prozesses bestimmen. Durch Verbinden der nicht gefrorenen Knoten untereinander erhält man jeweils das reduzierte nicht gefrorene Teilnetzwerk aus dem entsprechenden Netzwerkensemble, das im Folgenden untersucht wird.

Für eine Netzwerkrealisierung erhält man am Ende des stochastischen Prozesses N nicht gefrorene Knoten, von denen N_r (N_{c_2}) Knoten zwei nicht gefrorene Eingänge und eine reversible (kanalisierende \mathcal{C}_2) Knotenfunktion haben. Die anderen nicht gefrorenen Knoten haben (effektiv) einen Eingang.

Die Anwesenheit der Knoten mit zwei nicht gefrorenen Eingängen wird im Folgenden, siehe (4.1), (4.26) und (4.11), durch die Skalenvariable

$$a = \frac{N_r + N_{c_2}}{\sqrt{N}} = (1 + \gamma/\beta)y^{3/2} \quad (4.32)$$

beschrieben. Ihre Wahrscheinlichkeitsverteilung $f(a)$ folgt aus $f(a)da = G(y)dy$:

$$f(a) = \frac{2}{3a^{1/3}(1 + \gamma/\beta)^{2/3}} G\left(\left(\frac{a}{1 + \gamma/\beta}\right)^{2/3}\right). \quad (4.33)$$

Wie $G(y)$ wird $f(a)$ erst im thermodynamischen Limes $N^{ini} \rightarrow \infty$ exakt.

Nur ein kleiner Anteil der Knoten des nicht gefrorenen Teilnetzwerks bestimmt die Dynamik im Netzwerk. Diese relevanten Knoten werden wir mit Hilfe eines stochastischen Prozesses identifizieren, der die aufgrund ihrer Verbindungen im Netzwerk irrelevanten Knoten nacheinander findet und entfernt.

²In diesem Fall ist $\beta = 0$. Dennoch gelten nach Abschn. 4.2 alle folgenden Formeln und Aussagen mit einem effektiven Modellparameter $1 + (\gamma/\beta)_{eff}$, das bei Bedarf einfach bestimmt werden kann.

4. Kritische RBN mit $K = 2$

Die Knoten, die keine Ausgänge haben, sind trivialerweise irrelevant. Sie sind die Blätter der Bäume des Teilnetzwerks aus den nicht gefrorenen Knoten. Wir bestimmen zuerst ihre Anzahl in einer Netzwerkrealisierung.

Jeder der N nicht gefrorenen Knoten wählt seine nicht gefrorenen Eingangsknoten zufällig aus. Es stehen insgesamt $N(1 + a/\sqrt{N})$ Ausgänge zur Verfügung, da die Anzahl der Eingänge gleich der Anzahl der Ausgänge ist. Die Anzahl der Ausgänge eines Knotens ist nach Definition des Netzwerkensembles Poisson - verteilt mit dem Mittelwert $(1 + a/\sqrt{N})$. Der Anteil der Knoten, die keine Ausgänge haben, beträgt $\exp(-1 - a/\sqrt{N})$.

Als Nächstes werden alle Knoten mit Ausgängen in einen "beschrifteten" Container platziert. Ihre Anzahl wird mit N_b bezeichnet. Im Container befinden sich unter anderem irrelevante Knoten, deren Ausgänge ausschließlich mit anderen irrelevanten Knoten wie denjenigen außerhalb des Containers verbunden sind. Im Laufe des zu beschreibenden stochastischen Prozesses verringert sich N_b , da die irrelevanten Knoten aus dem Container entfernt werden, bis nur die relevanten übrig bleiben.

Die Gesamtanzahl der Aus- und Eingänge im beschrifteten Container beträgt anfangs jeweils $N(1 + a/\sqrt{N})$ und, durchschnittlich, $N(1 + a/\sqrt{N})(1 - \exp(-1 - a/\sqrt{N}))$. Es gibt anfangs mehr Ausgänge als Eingänge im Container. Die Eingänge eines beliebig ausgewählten Knotens außerhalb des Containers werden nun zufällig mit den Ausgängen im Container verbunden. Dieser verbundene ausgewählte Knoten wird von der weiteren Betrachtung ausgeschlossen. Die gewählten Ausgänge werden abgeschnitten, gegebenenfalls dabei entstehende Knoten ohne Ausgänge werden aus dem Container entfernt und außerhalb des Containers platziert. Solche Knoten und die übrigen Knoten außerhalb des Containers werden wie beschrieben nacheinander verbunden. Am Ende bleiben keine unverbundenen Knoten außerhalb des Containers und nur die relevanten Knoten im Container.

Da die Gesamtzahl der nicht verbundenen Eingänge im und außerhalb des Containers während der obigen Prozedur derjenigen der Ausgänge im Container gleich ist, haben beschriftete Knoten so lange mehr Ausgänge als Eingänge, bis alle Knoten außerhalb des Containers verbunden sind. Danach gibt es gleich viele Ein- und Ausgänge im Container.

Mit anderen Worten werden im beschriebenen Prozess einfach nacheinander zufällig ausgewählte Ausgänge im Container abgeschnitten, und die Knoten ohne Ausgänge werden aus dem Container entfernt. Dieser Prozess wird so lange fortgesetzt, bis die Anzahl der Ein- und Ausgänge im Container gleich werden. Die $N_b^{final} \equiv N_{rel}$ verbleibenden beschrifteten Knoten sind die relevanten Knoten.

Um die relevanten Knoten analytisch zu beschreiben ist es nützlich, den Prozess der Entfernung der Ausgänge bei $N_b^{final} \equiv N_{rel}$ zuerst weiter laufen zu lassen, bis keine Knoten mehr beschriftet sind, und dann rückwärts abzuspielen. Der Umkehrprozess entspricht der Richtung von unten nach oben im Beispiel in Abb. 4.4. Im Umkehrprozess kann man die ursprüngliche Poisson - Verteilung der Anzahl der Ausgänge wiederherstellen.

Der Umkehrprozess wird wie folgt zur analytischen Beschreibung herangezogen. Man startet mit N Knoten ohne Ausgänge und einem leeren beschrifteten Contai-

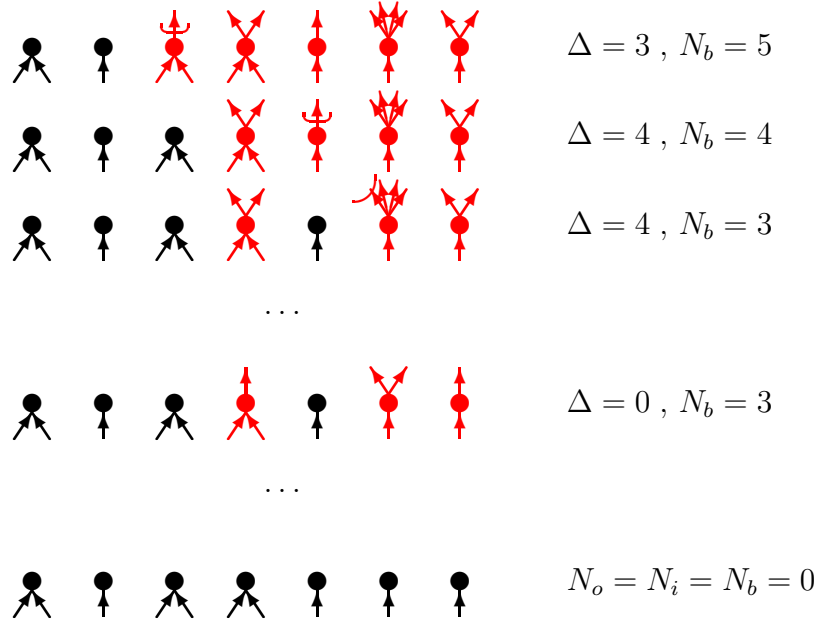


Abb. 4.4.: Schematische Darstellung des im Haupttext beschriebenen stochastischen Prozesses am Beispiel $N = 7$, $a = 3/\sqrt{7}$. Von oben nach unten werden die mit ovalen Bögen angedeuteten Ausgänge der Knoten im Container entfernt, von unten nach oben — hinzugefügt. N_b ist die Anzahl der Knoten im rot beschrifteten Container, $N_b^{final} \equiv N_{rel} = 3$. Δ ist der Überschuss an Ausgängen im Vergleich zu den Eingängen im Container. Die oberen drei Zeilen zeigen, dass Δ steigen, konstant bleiben und fallen kann.

ner. Die meisten der N Knoten haben einen Eingang. Der Anteil der Knoten mit zwei Eingängen beträgt a/\sqrt{N} . Man fügt nacheinander neue Ausgänge, einen pro einen Schritt, zufällig einem der N Knoten hinzu und beschriftet (verschiebt in den Container) die Knoten mit Ausgängen. Der Überschuss an den Ausgängen im Vergleich zu den Eingängen im Container ist zunächst Null und steigt in stochastischen Schritten nicht monoton an. Die beschrifteten Knoten sind die relevanten Knoten, wenn der Überschuss zum letzten Mal gleich Null ist.

Wir bezeichnen mit N_o die Anzahl der Ausgänge im Container minus die Anzahl der Knoten im Container und mit N_i die Anzahl der Knoten mit zwei Eingängen im Container. Der angesprochene Überschuss ist genau dann Null, wenn $N_o = N_i$.

Nur wenn ein Knoten im Container einen neuen Ausgang bekommt, kann N_o wachsen. Das passiert mit Wahrscheinlichkeit N_b/N . N_i erhöht sich um Eins mit Wahrscheinlichkeit a/\sqrt{N} . Für genügend große N kann N_b die Schritte zählen, sodass man für die mittleren Wachstumsraten von N_o und N_i erhält:

$$\left\langle \frac{dN_o}{dN_b} \right\rangle = \frac{N_b}{N}$$

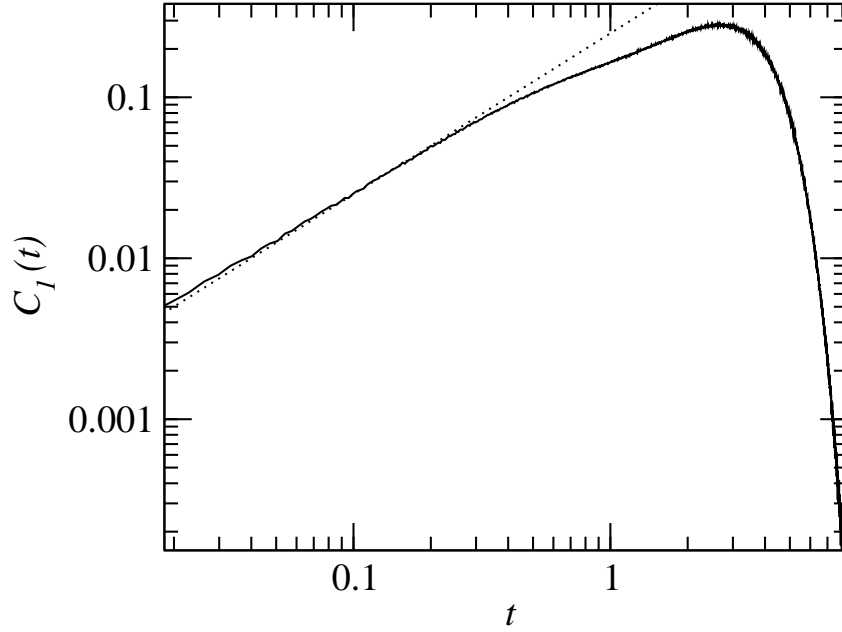


Abb. 4.5.: Die Funktion $C_1(t)$ aus numerischen Simulationen des entsprechenden stochastischen Prozesses. Die punktierte Gerade $0.25t$ stellt eine gute Näherung für kleine t dar.

und

$$\left\langle \frac{dN_i}{dN_b} \right\rangle = \frac{a}{\sqrt{N}}.$$

Durch die Einführung der “Zeitvariablen” t ,

$$t = \frac{N_b}{\sqrt{N}}, \quad (4.34)$$

lassen sich die Gleichungen für die mittleren Wachstumsraten vereinfachen:

$$\left\langle \frac{dN_o}{dt} \right\rangle = t$$

und

$$\left\langle \frac{dN_i}{dt} \right\rangle = a.$$

Im nächsten Abschnitt wird argumentiert, dass diese Vereinfachung mit einer Skalierung der Anzahl relevanter Knoten mit \sqrt{N} zusammenhängt.

Schließlich erhält man die Wahrscheinlichkeitsverteilungen für N_o und N_i :

$$P_o(N_o|t) = \frac{1}{N_o!} e^{-t^2/2} \left(\frac{t^2}{2} \right)^{N_o}, \quad (4.35)$$

$$P_i(N_i|t) = \frac{1}{N_i!} e^{-at} (at)^{N_i}. \quad (4.36)$$

Der stochastische Umkehrprozess kann als eine durch diese Verteilungen charakterisierte Zufallsbewegung der Größe $(N_o - N_i)$ angesehen werden, die im Mittel mit der Geschwindigkeit t nach rechts und mit der Geschwindigkeit a nach links schreitet. Die Zufallsbewegung wird beendet, wenn sie das letzte Mal den Ursprung verlässt. Die Wahrscheinlichkeitsverteilungen $\mathcal{C}_a(t)$ für die Zeit t , zu der dies passiert wurden numerisch berechnet, siehe Abb. 4.5 für den Fall $a = 1$. Bei kleinen t wächst die Wahrscheinlichkeitsverteilung linear mit t , da die Wahrscheinlichkeit eines Schrittes nach rechts dabei proportional zu t ist. Für $a = 0$ ergibt sich analytisch

$$\mathcal{C}_0(t) = -\frac{\partial P_o(0, t)}{\partial t} = te^{-t^2/2}. \quad (4.37)$$

4.3.2. Skalierung der Anzahl relevanter Knoten

Im Abschn. 4.3.1 konnte der stochastische Umkehrprozess durch Angabe von a vollständig beschrieben werden. Da die Anzahl der relevanten Knoten wegen $N_{rel} \equiv N_b^{final}$ und (4.34) aus dem Wert von t am Ende des Prozesses folgt, ergibt sich für die Wahrscheinlichkeitsverteilung der Anzahl der relevanten Knoten:

$$p_a(N_{rel})dN_{rel} = \mathcal{C}_a\left(N_{rel}/\sqrt{N}\right)dN_{rel}/\sqrt{N}. \quad (4.38)$$

Mit (4.11) und (4.32) folgt

$$\sqrt{N} = a^{1/3} \left(\frac{N^{ini}}{\beta + \gamma} \right)^{1/3},$$

und mit der Wahrscheinlichkeitsverteilung (4.33) für a ergibt sich für die relevanten Knoten

$$p(N_{rel}) = \int_0^\infty da f(a) \mathcal{C}_a \left(\frac{N_{rel} a^{-1/3}}{[N^{ini}/(\beta + \gamma)]^{1/3}} \right) \left(\frac{\beta + \gamma}{aN^{ini}} \right)^{1/3}. \quad (4.39)$$

Die von der Netzwerkgröße abhängige obere Integrationsgrenze kann für große Netzwerke und im Limes $N^{ini} \rightarrow \infty$ durch ∞ ersetzt werden. Benutzt man nun die SkalenvARIABLE

$$z = \frac{N_{rel}}{\left(\frac{N^{ini}}{\beta + \gamma} \right)^{1/3}} \quad (4.40)$$

und ihre Wahrscheinlichkeitsverteilung $P(z)$, so kommt für große N^{ini} das universelle Skalenverhalten der Anzahl der relevanten Knoten zum Vorschein:

$$P(z) = \int_0^\infty da \frac{f(a)}{a^{1/3}} \mathcal{C}_a \left(\frac{z}{a^{1/3}} \right). \quad (4.41)$$

Die Funktion $P(z)$ lässt sich effizient numerisch bestimmen. Dazu kombiniert man den stochastischen Prozess zur Bestimmung der Anzahl der nicht gefrorenen

4. Kritische RBN mit $K = 2$

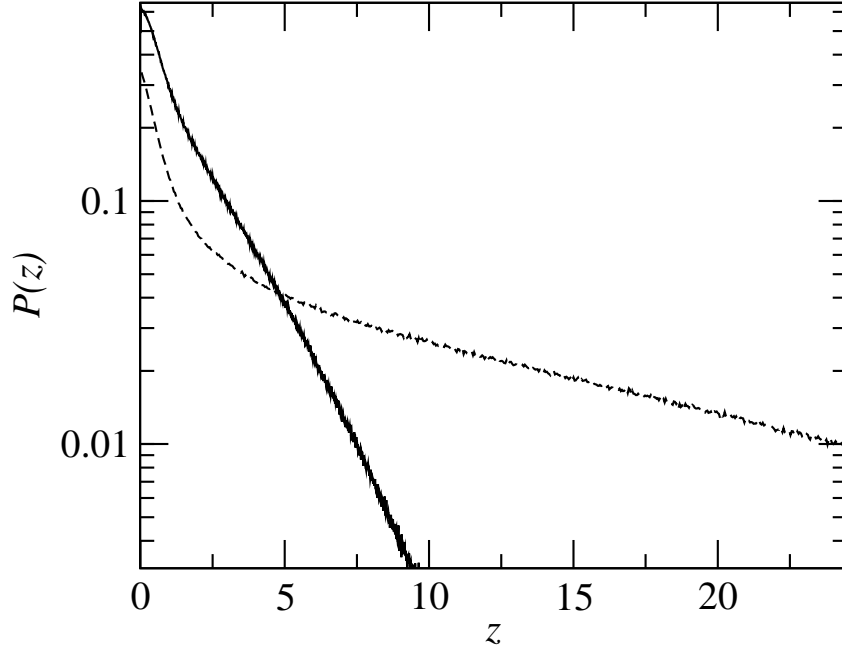


Abb. 4.6.: Die Funktion $P(z)$ für $\gamma/\beta = 0$ (durchgezogene Linie) und $\gamma/\beta = 4$ (gestrichelte Linie). Die Daten wurden über 10^7 verschiedene Simulationen der zwei gekoppelten stochastischen Prozesse aus dem Abschn. 4.1.3 und 4.3.1 gemittelt.

Knoten aus dem Abschn. 4.1.3 mit dem stochastischen Umkehrprozess aus dem Abschn. 4.3.1. Der erste gibt einen Wert von a zurück, im zweiten wird der Zeitpunkt bestimmt, zu dem die entsprechende Zufallsbewegung den Ursprung endgültig verlässt. Daraus gewinnt man einen Wert von z . Solche Simulationen werden dann statistisch ausgewertet, wobei eine möglichst große Stichprobe wichtig ist. Wegen (4.33) hängen die Ergebnisse ausschließlich von der Kombination der Modellparameter γ/β ab. In Abb. 4.6 ist $P(z)$ für $\gamma/\beta = 0$ und $\gamma/\beta = 4$ dargestellt, der Fall $\gamma/\beta = 4$ entspricht sRBN mit der gleichen Wahrscheinlichkeit für jede Knotenfunktion. Für $z \rightarrow 0$ kann ein nützliches mit den Daten vereinbares Ergebnis analytisch gezeigt werden:

$$\lim_{z \rightarrow 0} P(z) = \sqrt{2\pi}/[4(1 + \gamma/\beta)^{1/3}] \equiv \sqrt{2\pi}/(4\kappa^{1/3}). \quad (4.42)$$

Die mittlere Anzahl der relevanten Knoten ergibt sich zu

$$\bar{N}_{rel} = \int_0^\infty N_{rel} p(N_{rel}) dN_{rel} = \left(\frac{N^{ini}}{\beta + \gamma} \right)^{1/3} \int_0^\infty z P(z) dz. \quad (4.43)$$

\bar{N}_{rel} skaliert mit $(N^{ini})^{1/3}$.

4.3.3. Relevante Knoten mit zwei relevanten Eingängen

In diesem Abschnitt wird gezeigt, dass die Anzahl m der relevanten Knoten mit zwei relevanten Eingängen im thermodynamischen Limes endlich ist. Außerdem stelle ich einige Formeln für die weitere Untersuchung der relevanten Knoten und ihrer Organisation in Komponenten bereit.

Die Wahrscheinlichkeit dafür, dass die Anzahl der relevanten Knoten (4.40), von denen m Knoten zwei relevante Eingänge haben, im Intervall $[N_{rel}(z), N_{rel}(z + dz)]$ liegt, wird mit $\tilde{P}(m; z)dz$ bezeichnet. Aus den Gleichungen (4.41), (4.35), (4.36) und (4.34) folgt

$$\begin{aligned} \tilde{P}(m; z) &= \int_0^\infty da \frac{f(a)}{a^{1/3}} \mathcal{C}_a \left(\frac{z}{a^{1/3}} \right) \\ &\quad \times \frac{P_o(m|za^{-1/3}) P_i(m|za^{-1/3})}{\sum_l P_o(l|za^{-1/3}) P_i(l|za^{-1/3})}. \end{aligned} \quad (4.44)$$

Der letzte Bruch in dieser Formel stellt sicher, dass die Anzahl der Knoten mit zwei relevanten Eingängen am Ende des stochastischen Umkehrprozesses m beträgt, wobei aus dem Abschn. 4.3.1 $N_o = N_i = m$ gilt.

Die Wahrscheinlichkeitsverteilung für die Anzahl m der relevanten Knoten mit zwei relevanten Eingängen ergibt sich zu

$$\begin{aligned} p(m) &= \int_0^\infty \tilde{P}(m; z) dz \\ &= \frac{1}{(m!)^2} \int_0^\infty dz \frac{(z^3/2)^m P(z)}{I_0(\sqrt{2z^3})}, \end{aligned} \quad (4.45)$$

wobei $I_n(z)$ die modifizierten Bessel - Funktionen der ersten Art bezeichnet. Es gilt $I_0(x) = \sum_{m=0}^\infty (\frac{x^m}{2^m m!})^2$, $I_1(x) = 2/x \sum_{m=0}^\infty m (\frac{x^m}{2^m m!})^2$ für $n = 0$ und $n = 1$.

Für die mittlere Anzahl der Knoten mit zwei relevanten Eingängen \bar{m} erhält man

$$\bar{m} = \frac{1}{2} \int_0^\infty dz \frac{\sqrt{2z^3} I_1(\sqrt{2z^3})}{I_0(\sqrt{2z^3})} P(z) dz. \quad (4.46)$$

Das Integral konvergiert gegen einen endlichen Wert, da $P(z)$ für große z exponentiell abfällt, siehe Abb. 4.6, und der Rest des Integranden schneller als linear, jedoch langsamer als quadratisch in z ansteigt. Dass \bar{m} endlich und für große Netzwerke von deren Größe unabhängig ist, geht auf die exponentiell abfallende Abhängigkeit der Funktionen P_o und P_i von m in (4.45) zurück.

Die Abb. 4.7 zeigt die Ergebnisse von einigen direkten Computersimulationen zur Bestimmung von \bar{m} . Mit steigenden N^{ini} streben die Kurven gegen einen endlichen Wert.

Die Abhängigkeit der funktionalen Form $p(m)$ (und folglich auch \bar{m}) von den Modellparametern β und γ wird durch die Abhängigkeit der funktionalen Form

4. Kritische RBN mit $K = 2$

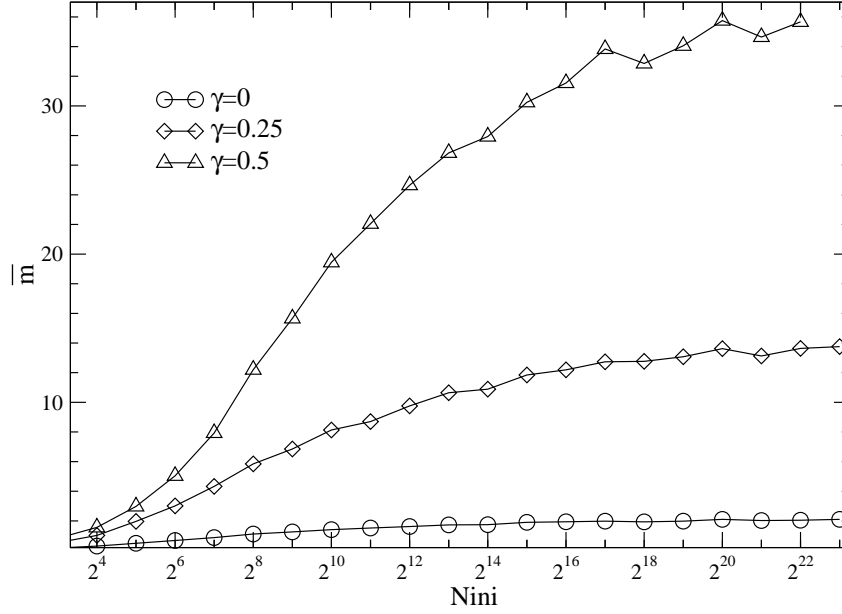


Abb. 4.7.: Mittlere Anzahl der Knoten mit zwei relevanten Eingängen für $\beta = 1/8$ als Funktion der Netzwerkgröße N^{ini} . Die Mittelung erfolgte über 10^4 Netzwerkrealisierungen.

$P(z)$ von $\kappa \equiv 1 + \gamma/\beta$ bestimmt. $P(z)$ wird für größere κ breiter und nimmt daher kleinere Werte nur bei kleinen z in (4.46) an. Folglich wächst \bar{m} mit κ . Bei konstanter Netzwerkgröße führen größere κ also zu mehr relevanten Knoten mit zwei relevanten Eingängen.

4.4. Analogie zwischen $K = 2$ und $K = 1$ Netzwerken

Im Folgenden möchte ich untersuchen, inwieweit die kritischen $K = 2$ und $K = 1$ Netzwerke vergleichbar sind. Als eine Anwendung der Ergebnisse wird die mittlere Anzahl der Knoten in den stabilisierenden Schleifen (siehe Abschn. 4.2.1) der kanalisierenden $K = 2$ Netzwerke bestimmt.

Kritische $K = 1$ Netzwerke mit N Knoten entsprechen den nicht gefrorenen Teilnetzwerken der kritischen $K = 2$ Netzwerke mit dem Anteil $a = 0$ der Knoten mit zwei nicht gefrorenen Eingängen. Die Wahrscheinlichkeitsverteilung für die Anzahl der relevanten Knoten ergibt sich mit (4.37) als ein Spezialfall von (4.38) zu:

$$p_0(N_{rel}) = \frac{1}{\sqrt{N}} \mathcal{C}_0 \left(\frac{N_{rel}}{\sqrt{N}} \right) = \frac{N_{rel}}{N} e^{-N_{rel}^2/2N}. \quad (4.47)$$

Die mittlere Anzahl der relevanten Knoten ist für große Netzwerke proportional zu \sqrt{N} . Wenn die Ein- und Ausgänge der relevanten Knoten untereinander verbunden werden, ergeben sich die relevanten Komponenten, im Fall $a = 0$ — einfache

4.4. Analogie zwischen $K = 2$ und $K = 1$ Netzwerken

Schleifen von Knoten. Beispielsweise aus [Dros 05b] ist bekannt, dass die mittlere Anzahl dieser Schleifen $C \approx \ln \sqrt{N}$ beträgt, genauer aus [Samu 05]

$$C \simeq 1/2(\ln(2N) + \gamma_E), \quad (4.48)$$

$\gamma_E \approx 0.577$ ist die Euler - Mascheroni - Konstante.

Die Anzahl der einfachen Schleifen der Länge l ist in kritischen $K = 1$ Netzwerken Poisson - verteilt, sie kommen im Grenzfall $N_{rel} \rightarrow \infty$ vor mit der mittleren Wahrscheinlichkeit

$$P(l) = 1/l \text{ für } l < l_c \text{ mit } l_c \sim \sqrt{N}. \quad (4.49)$$

Das folgt für $K = 1$ Netzwerke anschaulich aus einer expliziten Konstruktion des relevanten Teilnetzwerks aus N_{rel} Knoten. Wir starten mit einem relevanten Knoten und wählen seinen relevanten Eingangsknoten zufällig aus. Anschließend wird der Eingangsknoten des ersten Eingangsknotens usw. ausgewählt, bis sich die Schleife am Startknoten schließt. Für kleine Schleifengrößen beträgt die Wahrscheinlichkeit für die Schließung nach dem Hinzufügen des l -ten Knotens $1/N_{rel}$. Die Wahrscheinlichkeit dafür, dass ein Knoten in einer Schleife der Länge l sitzt, ist damit $1/N_{rel}$, die mittlere Anzahl der Knoten in den Schleifen der Länge l ist 1 und die Anzahl der Schleifen der Länge l ist Poisson - verteilt mit dem Mittelwert $1/l$ für genügend kleine l .

In kritischen $K = 2$ Netzwerken haben nach Abschn. 4.3.2 und 4.3.3 die meisten der $N_{rel} \sim (N^{ini})^{1/3}$ relevanten Knoten effektiv einen Eingang, der Anteil der endlichen Anzahl der Knoten mit zwei relevanten Eingängen verschwindet im Limes $N^{ini} \rightarrow \infty$. Eine Analogie zu den relevanten $K = 1$ Teilnetzwerken ist daher nahe liegend. Die relevanten Komponenten werden wieder anschaulich durch Verbinden der Ein- und Ausgänge der relevanten Knoten konstruiert. Eine kleine Komponente besteht im betrachteten Grenzfall $N^{ini} \rightarrow \infty$ fast immer aus Knoten mit einem relevanten Eingang. Analog zum $K = 1$ Fall erhält man eine Poisson - Verteilung für die Anzahl der einfachen Schleifen der Länge l , sie kommen vor jeweils mit der Wahrscheinlichkeit

$$P(l) = 1/l \text{ für } l < l_c \text{ mit } l_c \sim (N^{ini})^{1/3}. \quad (4.50)$$

Die Gesamtzahl der relevanten Knoten in den Schleifen beträgt ϵN_{rel} mit $\epsilon \lesssim 1$. Die übrigen $(1 - \epsilon)N_{rel}$ relevanten Knoten sitzen in komplexen relevanten Komponenten mit Knoten mit zwei relevanten Eingängen. Mit der endlichen Anzahl der komplexen Komponenten befassen wir uns ausführlich im Abschn. 4.5.

Abschließend kann die mittlere Größe der nicht gefrorenen Bäume, die in relevanten Komponenten wurzeln, abgeschätzt werden. Dazu teilt man einfach die mittlere Anzahl der nicht gefrorenen Knoten durch die mittlere Anzahl der relevanten Knoten. Für die mittlere Baumgröße ergibt sich größenordnungsmäßig $N^{1/3}$ in kritischen $K = 2$ Netzwerken und $N^{1/2}$ in kritischen $K = 1$ Netzwerken.

4.4.1. Größe stabilisierender Schleifen

Mit den im letzten Abschnitt zusammengestellten Ergebnissen zu den relevanten Komponenten kann man im Fall der kanalisierenden \mathcal{C}_2 Netzwerke die mittlere Anzahl der Knoten in den einfachen stabilisierenden Schleifen abschätzen sowie einige bereits bekannte Ergebnisse bestätigen, vgl. Abschn. 4.2.2.

Ich betrachte ein kanalisierendes \mathcal{C}_2 Netzwerk mit N Knoten. Die Wahrscheinlichkeit, dass ein Knoten eine Selbstverbindung hat und auf seinem Mehrheitsbit sich selbst kanalisiert (eine stabilisierende Schleife bildet) beträgt $2/N \times 1/2 = 1/N$. Man erhält im Mittel eine einfache stabilisierende Schleife der Länge 1 pro Netzwerk. Die mittlere Anzahl der einfachen stabilisierenden Schleifen der Länge 2 pro Netzwerk beträgt analog $[(\frac{N}{2})(2/N)^2] \times (\frac{1}{2})^2 \approx \frac{1}{2}$. Für einfache stabilisierende Schleifen der Länge $l > 2$ muss man mit einem weiteren Faktor die Anzahl der Möglichkeiten berücksichtigen, aus l Knoten eine gerichtete Schleife zu bauen. Die Anzahl der einfachen stabilisierenden Schleifen der Länge l ergibt sich zu $1/l$. Die Gesamtzahl der Knoten in den stabilisierenden Schleifen f_0 ist

$$f_0 = \sum_{l=1}^{l_c} \frac{1}{l} l = l_c, \quad (4.51)$$

mit der Abschneidelänge l_c . Die einfache probabilistische Betrachtung ist gültig für $l_c \ll N$.

Die Größe l_c zu berechnen wird uns die Analogie zu einem anderen Problem mit bekannter Lösung helfen. Wir wollen nämlich zuerst alle relevanten Knoten eines kritischen Netzwerks mit N Knoten, $K = 2$ und Knotenfunktionen aus den Klassen \mathcal{R} und \mathcal{F} (also ohne kanalisierende Funktionen) identifizieren. Nach Abschn. 4.4 sind die meisten relevanten Komponenten einfache Schleifen. Wir fragen, ob ein ausgezeichnete Knoten in einer relevanten Schleife sitzt. Folgende Fälle sind zu unterscheiden.

- Mit Wahrscheinlichkeit $1/4$ haben beide Eingangsknoten des ausgezeichneten Knotens gefrorene Funktionen und der Knoten ist irrelevant.
- Mit Wahrscheinlichkeit $1/2$ haben die Eingangsknoten des ausgezeichneten Knotens eine gefrorene und eine reversible Funktion. In diesem Fall wird die Verbindung zum Eingangsknoten mit der reversiblen Funktion markiert und der Eingangsknoten wird für die Untersuchung im nächsten Schritt vorgemerkt.
- Mit Wahrscheinlichkeit $1/4$ haben beide Eingangsknoten reversible Funktionen, beide Verbindungen werden markiert, beide Eingangsknoten werden vorgemerkt.

Diese Prozedur wird für die vorgemerkten Knoten iterativ fortgesetzt, bis eine markierte Verbindung zum ursprünglich ausgezeichneten Knoten führt (sodass er in einer relevanten Schleife sitzt) oder bis keine neuen Verbindungen markiert werden

können (sodass der ausgezeichnete Knoten zu keiner relevanten Schleife gehört). Nach Abschn. 4.4 gibt es im betrachteten Netzwerk im Mittel $1/l$ einfache relevante Schleifen der Länge $l < l_c$ mit der Abschneidelänge l_c , die mit $N^{1/3}$ skaliert. Die einfachen relevanten Schleifen wie übrigens auch die wenigen komplexen relevanten Komponenten beinhalten größenordnungsmäßig $N^{1/3}$ Knoten. Der ausgezeichnete Knoten sitzt in einer relevanten Schleife mit der Wahrscheinlichkeit $N^{1/3}/N$.

Als Nächstes wird gefragt, welche Knoten in \mathcal{C}_2 Netzwerken in den stabilisierenden Schleifen sitzen. Die Situation ist analog zu derjenigen im obigen Problem, wobei die Knoten in den stabilisierenden Schleifen hier den Knoten in den relevanten Schleifen oben entsprechen. Mit Wahrscheinlichkeit $1/4$ kanalisiert der Mehrheitsbit keiner der Eingangsknoten einen ausgezeichneten Knoten, der ausgezeichnete Knoten befindet sich demnach in keiner stabilisierenden Schleife. Mit Wahrscheinlichkeit $1/2$ kanalisiert der Mehrheitsbit von nur einem der Eingänge den ausgezeichneten Knoten, der Eingangsknoten wird vorgemerkt. Mit Wahrscheinlichkeit $1/4$ kanalisiert der Mehrheitsbit beider Eingänge den ausgezeichneten Knoten, beide Eingangsknoten werden vorgemerkt. Die vorgemerkten Knoten werden auf die gleiche Weise nacheinander untersucht, bis der ursprünglich ausgezeichnete Knoten vorgemerkt wird und sich damit in einer stabilisierenden Schleife befindet oder bis keine neuen Knoten vorgemerkt werden können und der ausgezeichnete Knoten damit in keiner stabilisierenden Schleife sitzt. Aus der Analogie zum obigen Problem folgt, dass es im Mittel $1/l$ einfache stabilisierende Schleifen der Länge $l < l_c$ gibt. Des Weiteren skaliert die Abschneidelänge l_c mit $N^{1/3}$. Die Gesamtzahl der Knoten in einfachen sowie in allgemeinen stabilisierenden Schleifen ist proportional zu $N^{1/3}$, vgl. Abschn. 4.2.2.

4.5. Relevante Komponenten

Die folgenden Ergebnisse wurden in [Kauf 06] veröffentlicht. Im Abschn. 4.5.1 wird die Wahrscheinlichkeitsverteilung der komplexen relevanten Komponenten mit einer festen Anzahl der Knoten mit zwei relevanten Eingängen diskutiert. Sie wird im Limes großer Netzwerke unabhängig von der Größe der Netzwerke. Phänomenologische und analytische Überlegungen im Abschn. 4.5.2 führen zu folgenden Ergebnissen. Es wird bestätigt, dass die Anzahl der relevanten Komponenten logarithmisch mit der Netzwerkgröße wächst. Es wird gezeigt, dass ein größerer Anteil der nicht gefrorenen Knoten mit zwei relevanten Eingängen zu einer kleineren Anzahl der relevanten Komponenten und einer größeren und zunehmend komplexen größten relevanten Komponente führt. Die Anzahl der relevanten Komponenten reduziert sich dabei annähernd um die mittlere Größe des Logarithmus der Anzahl der relevanten Knoten mit zwei relevanten Eingängen. Alle Ergebnisse werden durch numerische Simulationen bestätigt.

4. Kritische RBN mit $K = 2$

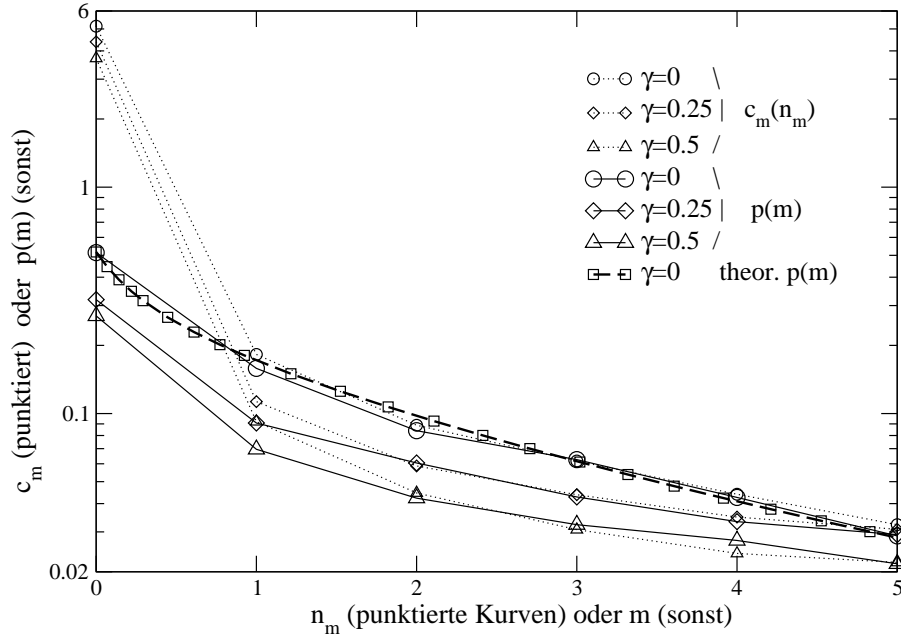


Abb. 4.8.: Numerische Simulation der Anzahl $c_m(n_m)$ der Komponenten pro Netzwerk mit n_m Knoten mit zwei relevanten Eingängen (punkturierte Kurven). Numerische Simulation der Wahrscheinlichkeitsverteilung $p(m)$ der Anzahl m der Knoten mit zwei relevanten Eingängen pro Netzwerk (durchgezogene Kurven). Die theoretische Vorhersage für die Wahrscheinlichkeitsverteilung $p(m)$ (gestrichelte Kurve). Benutzte Modellparameter: $\beta = 1/8$ und $\gamma = 0$ (O, \square), $\gamma = 1/4$ (\diamond), $\gamma = 1/2$ (\triangle). Alle Simulationen wurden über 8000 Netzwerkrealisierungen gemittelt. Die Netzwerkgröße war $N \sim 10^7$ (punkturierte Kurven), $N \sim 10^5$ (durchgezogene Kurven). Die Näherung $P_{\gamma=0}^{fit}(z) = 0.62 e^{-0.65z}$ wurde benutzt um die gestrichelte Kurve zu erhalten.

4.5.1. Verteilung der komplexen relevanten Komponenten

Nachdem die einfachen relevanten Komponenten in kritischen RBN mit N Knoten und $K = 2$ im Abschn. 4.4 untersucht wurden, beschäftige ich mich hier mit komplexen relevanten Komponenten, die Knoten mit zwei relevanten Eingängen beinhalten. Mich interessiert zunächst, wie häufig die Komponenten mit einer vorgegebenen Anzahl der Knoten mit zwei relevanten Eingängen im Netzwerkensemble vorkommen.

Mit der in (4.45) angegebenen Wahrscheinlichkeit $p(m)$ enthält das relevante Teilnetzwerk m Knoten mit zwei relevanten Eingängen. Bei einem festgehaltenen Wert von m kommen alle topologisch möglichen relevanten Teilnetzwerke mit derselben Wahrscheinlichkeit vor. Sie bilden das Netzwerkensemble $\mathcal{E}_{N,m}$. Für dieses Ensemble wurde im Abschn. 3.2.2 gezeigt, dass in 90% der Fälle oder häufiger alle Knoten mit zwei relevanten Eingängen in derselben Komponente liegen. Für ein gegebenes Netzwerk folgt daraus, dass eine komplexe Komponente mit m Knoten mit

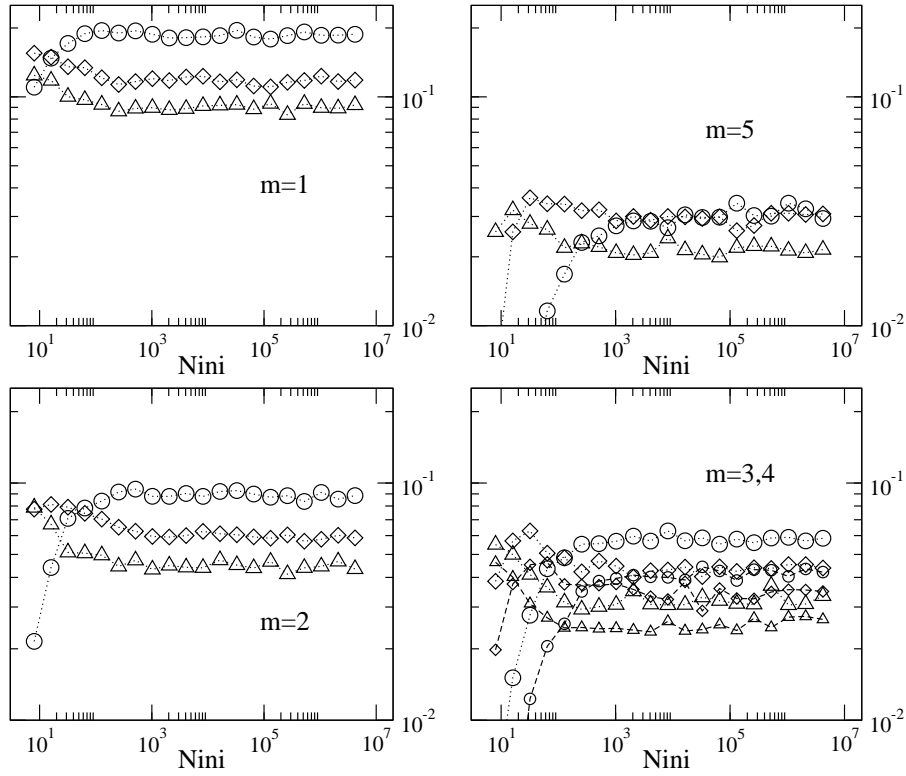


Abb. 4.9.: Die Anzahl der relevanten komplexen Komponenten mit m Knoten mit zwei relevanten Eingängen ist endlich für große Netzwerke. Die Symbole O, \diamond und \triangle entsprechen den Modellparametern $\beta = 1/8$ und $\gamma = 0, 1/4$ und $1/2$. Die Datenpunkte wurden über 8000 Netzwerkrealisierungen gemittelt.

zwei relevanten Eingängen bis auf wenige Ausnahmen genau dann auftritt, wenn die Anzahl der Knoten mit zwei relevanten Eingängen in diesem Netzwerk m ist. Die detailliertere Betrachtung weiter unten bestätigt unter anderem, dass $p(m)$ bei $m > 1$ tatsächlich die Wahrscheinlichkeit für das Auftreten einer Komponente mit m Knoten mit zwei relevanten Eingängen angibt.

Numerische Simulationsergebnisse sind in Abb. 4.8 dargestellt. Dazu wurden Netzwerkensembles künstlich mit Hilfe der stochastischen Prozesse aus den Abschn. 4.1.3 und 4.3.1 generiert. Die punktierten Kurven für die mittlere Anzahl c_m der Komponenten mit n_m Knoten mit zwei relevanten Eingängen pro Netzwerk wurden dann durch einfaches Abzählen für die verschiedenen Modellparameter (steigende $\kappa = 1 + \gamma/\beta$ entsprechen aus (4.46) steigenden \bar{m}) berechnet. Die durchgezogenen Kurven in Abb. 4.8 zeigen die direkt numerisch bestimmte Wahrscheinlichkeitsverteilung $p(m)$. Für $m > 1$ und $n_m > 1$ rührt die Abweichung der punktierten Linien von den entsprechenden durchgezogenen Linien im Wesentlichen von den statistischen Fluktuationen her. Sie werden größer bei größeren Netzwer-

4. Kritische RBN mit $K = 2$

ken, weswegen vergleichsweise kleine Netzwerke mit $N \sim 10^5$ benutzt wurden um die durchgezogenen Kurven zu erhalten.

In den seltenen Netzwerkrealisierungen mit zwei komplexen relevanten Komponenten hat die kleinere Komponente meistens einen relevanten Knoten mit zwei relevanten Eingängen ($m = 1$), daher der erhöhter Wert von c_m im Vergleich zu $p(m)$ bei $n_m = m = 1$. (Die Differenz bei $n_m = m = 1$ in Abb. 4.8 ist größer bei größeren \bar{m} , da dann die Verteilung $p(m)$ flacher ist und für eine größere Anzahl der m - Werte eine $m = 1$ Komponente sich abspalten kann.) Die gestrichelte Kurve zeigt schließlich die analytische Berechnung von $p(m)$ mit Hilfe der Formel (4.45), die durch eine nahezu perfekte Übereinstimmung mit den numerischen Daten die Formel bestätigt. Ich habe die theoretische Kurve nur für $\gamma = 0$ berechnet. Dass die Anzahl der Komponenten mit einer konstanten Anzahl m der Knoten mit zwei relevanten Eingängen im thermodynamischen Limes tatsächlich unabhängig von der Systemgröße wird, zeigt die Abb. 4.9.

Die Abb. 4.8 beinhaltet weitere nützliche Daten. Für $m = 0$ geben die Funktionswerte der durchgezogenen Kurven den Anteil der Netzwerke ohne Knoten mit zwei relevanten Eingängen an. Diese Anteile,

$$p(0) = 0.51, 0.32, 0.27 \text{ für } \gamma = 0, 0.25, 0.5, \quad (4.52)$$

brauchen wir später für die Abb. 4.10. Die Netzwerkrealisierungen ohne Knoten mit zwei relevanten Eingängen sind nicht von den entsprechenden Netzwerkrealisierungen im Modell mit $K = 1$ zu unterscheiden. Der Anteil der "gefrorenen" Netzwerke ohne relevante Knoten verschwindet übrigens wie $(\beta/N)^{1/3}$ im Limes $N \rightarrow \infty$, obwohl $\lim_{z \rightarrow 0} P(z) \neq 0$. Die gestrichelten Kurven geben bei $n_m = 0$ die Anzahl der einfachen Schleifen pro Netzwerk an,

$$c_m(0) = 5.1, 4.4, 3.7 \text{ für } \gamma = 0, 0.25, 0.5. \quad (4.53)$$

Diese Werte werden im nächsten Abschnitt (siehe Abb. 4.10) phänomenologisch bestätigt.

Wir haben den Zusammenhang der Funktion $p(m) \equiv p_\kappa(m)$ mit der Verteilungsfunktion für die Anzahl der topologischen Verzweigungen in den komplexen relevanten Komponenten untersucht. Abschließend wird die Abhängigkeit der Funktion $p_\kappa(m)$ vom Modellparameter κ genauer betrachtet.

Analog zur Diskussion am Ende des Abschn. 4.3.3 ist $p(m)$ für konstante κ monoton fallend. Die Funktion wird flacher für größere κ . Wie verhält sich nun $p_\kappa(m)$ als Funktion von κ für ein konstantes m ? Eine Analyse der Gleichung (4.45) zeigt, dass der andere neben $P(z)$ Faktor im Integranden ein globales Maximum in z hat, das sich zu größeren Werten von z bei steigendem m verschiebt. Die Funktion $P(z)$ selbst fällt bekanntlich monoton ab, sie wird für größere κ flacher und nimmt anfangs (bei z nahe 0) kleinere Werte an. Zusammenfassend, für konstante m hat die Funktion $p_\kappa(m)$ ein Maximum, dessen Lage sich auf der κ - Achse mit steigenden m zu größeren κ verschiebt. Man erkennt Hinweise auf dieses Verhalten in den Abb. 4.8 und 4.9, in denen bei $m > 4$ die Datenpunkte für $\kappa = 1$ nicht mehr

größere Werte als die entsprechenden Datenpunkte für andere κ annehmen. Für sehr große κ und nicht zu große m kann die Funktion $P(z)$ in (4.45) durch ihren Wert $P(z=0)$ aus (4.42) angenähert werden, sodass explizit $p_\kappa(m) \sim \kappa^{-1/3}$ gilt.

In wenigen Worten zeigt die Analyse der Verteilungsfunktion $p_\kappa(m)$, dass mit zunehmenden Werten von κ die komplexen relevanten Komponenten mit einer kleineren Anzahl m der Knoten mit zwei relevanten Eingängen seltener auftreten, während die komplexen Komponenten mit größeren $m \sim \bar{m}(\kappa)$ häufiger auftreten.

4.5.2. Anzahl der relevanten Komponenten

In diesem Abschnitt wird eine analytische Abschätzung des Einflusses der Knoten mit zwei relevanten Eingängen auf die Anzahl der relevanten Komponenten und auf die Anzahl der einfachen Schleifen, sowie numerische Simulationen zur Wahrscheinlichkeitsverteilung der Anzahl der relevanten Komponenten präsentiert.

Wir konstruieren die relevanten Komponenten aus den N_{nf} Knoten des nicht gefrorenen Teilnetzwerks eines $K = 2$ Netzwerks mit N Knoten. Zunächst wird ein Eingang von jedem Knoten mit zwei nicht gefrorenen Eingängen außer Acht gelassen, er wird später verbunden. Dann bleibt ein handliches kritisches $K = 1$ Netzwerk zurück. Es besteht aus einfachen Schleifen und aus den in ihnen wurzelnden Bäumen von Knoten. Die mittlere Anzahl der einfachen Schleifen beträgt nach (beispielsweise) [Samu 05] asymptotisch

$$C \simeq 1/2(\ln(2N_{nf}) + \gamma_E), \quad (4.54)$$

$\gamma_E \approx 0.577$ ist die Euler - Mascheroni - Konstante. Für die Breite der entsprechenden Verteilung gilt asymptotisch [Samu 05]

$$\sigma^2 = C - \pi^2/8. \quad (4.55)$$

Die Gleichungen (4.54) und (4.55) gelten für einen konstanten Wert von N_{nf} . Durch Einsetzen des Mittelwerts $\bar{N}_{nf} \approx 0.62(N/\beta)^{2/3}$ aus (4.14) erhält man $\bar{C} \approx 6.4$ für $N = 2^{23}$. Um genauer zu sein, muss man die volle Wahrscheinlichkeitsverteilung für die Anzahl der nicht gefrorenen Knoten N_{nf} und für die Anzahl der einfachen Schleifen C benutzen. Man erhält die mittlere Anzahl der einfachen Schleifen als eine Kombination der exakten Ergebnisse (II.C16)³ und (II.C17)⁴ aus [Samu 05] mit der umskalierten Wahrscheinlichkeitsverteilung $G(y)$ aus (4.15):

$$\begin{aligned} \bar{C}(N) &\simeq \sum_{\mu=1}^N \mu \sum_{L=1}^N \frac{1}{L!} \left[\begin{matrix} L \\ \mu \end{matrix} \right] \sum_{N_{nf}=1}^N (N/\beta)^{-2/3} \\ &G(N_{nf}(N/\beta)^{-2/3}) \binom{N_{nf}}{L} \frac{LL!}{N_{nf}^{L+1}}. \end{aligned} \quad (4.56)$$

³Die Verteilung für die Anzahl L der Knoten in den Schleifen, für ein konstantes N_{nf} : $P(L) = N_{nf}!L / N_{nf}^{L+1} / (N_{nf} - L)!$.

⁴Die volle Verteilung für L und für die Anzahl μ der Schleifen, für ein konstantes N_{nf} : $P(\mu, L) = \left[\begin{matrix} L \\ \mu \end{matrix} \right] / L!$.

4. Kritische RBN mit $K = 2$

Ich habe ein Computerprogramm in Mathematica zur effizienten Berechnung von (4.56) geschrieben, das für $N = 2^{23}$ den Wert $\bar{C} = 6.13$ ausgibt. Dieser Wert wird im Folgenden benutzt, indessen wir bemerken, dass die obige Rechnung mit dem Mittelwert eine gute Abschätzung des exakten Resultats darstellt.

Von den \bar{C} einfachen Schleifen wird nur ein Teil von den relevanten Knoten mit zwei relevanten Ein- oder Ausgängen “belegt”. Nach Abschn. 4.5.1 kann man annehmen, dass diese relevanten Knoten nach Verbinden der noch nicht verbundenen zweiten Eingänge in einer einzigen komplexen Komponente liegen werden, sodass man mit (4.56) auf die gesuchte mittlere Anzahl der relevanten Schleifen schließen kann, sobald die Anzahl der “belegten” Schleifen \bar{n} bekannt ist. Wir kommen zur Abschätzung von \bar{n} nach der Anmerkung, dass wir (wie im obigen Beispiel für \bar{C}) Mittelwerte von bestimmten Funktionen implizit durch die Funktionen der Mittelwerte der Argumente annähern werden.

Wegen (4.49) sind die mittleren Gesamtanzahlen der Schleifen mit Längen aus logarithmisch konstanten Intervallen alle gleich. Mit den Bezeichnungen l_1 und l_2 für die untere und die obere Grenze eines Intervalls gibt es im Mittel jeweils eine einfache Schleife pro Intervall der logarithmischen Größe $l_2/l_1 = e$. Wir betrachten solche Intervalle mit den oberen Grenzen l_2 gleich e, e^2 , usw. Für zwei benachbarte Intervalle beträgt das mittlere Quotient der Anzahl der Knoten in den Schleifen der zwei Intervalle e . Das mittlere Quotient der Gesamtanzahl der Knoten in den Komponenten aus den beiden Intervallen ist auch e , da für jedes Intervall die mittlere Gesamtanzahl der Knoten in den Schleifen und in den in ihnen wurzelnden Bäumen proportional zur mittleren Anzahl der Knoten in den Schleifen ist.

Mit diesen Annahmen beträgt der Anteil der Knoten in der s -ten Komponente, ab der größten Komponente gezählt, $\lambda e^{-(s-1)}$ mit $\lambda = 1 - 1/e \approx 0.632$. Dabei wurde $e^{-\bar{C}} \ll 1$ berücksichtigt. Der Wert $\lambda = 1 - 1/e$ liegt nah am numerisch bestimmten Wert $\lambda \approx 0.624$ für den mittleren asymptotischen Anteil der Knoten in der größten nicht gefrorenen Komponente des Netzwerkensembles, das aus den kritischen $K = 1$ Netzwerken mit einer beliebig großen, jedoch konstanten Anzahl der nicht gefrorenen Knoten besteht. Im Allgemeinen muss die Anzahl der nicht gefrorenen Knoten natürlich durch eine Wahrscheinlichkeitsverteilung charakterisiert werden. Dennoch haben wir eine zusätzliche Bestätigung der Plausibilität der obigen Abschätzungen.

In einem Netzwerk wird mit m die Anzahl der relevanten Knoten mit zwei relevanten Eingängen bezeichnet, einer der Eingänge wird wie gesagt als abgeschnitten angesehen. Die Anzahl der Knoten mit zwei relevanten Ein- oder Ausgängen beträgt in der größten nicht gefrorenen Komponente größenordnungsmäßig $2m\lambda$, die kleineren Komponenten beinhalten $2m\lambda/e, 2m\lambda/e^2, \dots, 1$ Knoten mit zwei relevanten Ein- oder Ausgängen. Um die angestrebte Abschätzung für \bar{n} zu erhalten wird im Folgenden vernachlässigt, dass die Verteilung der Anzahl der Knoten mit zwei relevanten Eingängen $p(m)$ ($p(m)$ wurde im Abschn. 4.5.1 diskutiert) von der Anzahl der relevanten Knoten abhängt und daher mit der Anzahl der nicht gefrorenen Knoten in den Schleifen und mit λ korreliert.

Die mittlere Gesamtzahl $n(m)$ der nicht gefrorenen Komponenten mit Knoten mit zwei relevanten Ein- oder Ausgängen (einer von ihnen ist jeweils abgeschnitten) er-

gibt sich für $m \geq 1$ aus der Normierungsbedingung $2m\lambda e^{-(n(m)-1)} = 1$. Für $m = 1$ erhält man insbesondere $n(1) \simeq 1.234$. Das exakte Resultat $n(1) = 4/3$ weicht von dieser Abschätzung um weniger als 10% ab. Um das exakte Resultat zu erklären verbinden wir den abgeschnittenen Eingang vom Knoten mit zwei relevanten Eingängen und betrachten alle möglichen auf diese Weise entstehenden komplexen relevanten Komponenten. Das sind (ungeachtet der Bäume) die zwei möglichen $m = 1$ Komponenten aus dem Abschn. 3.2.1, wobei die beiden Eingänge des Knotens mit zwei relevanten Eingängen mit der gleichen Wahrscheinlichkeit abgeschnitten worden sein können. Die Schleife mit einer Querverbindung (ggf. mit Bäumen) kommt nach Abschn. 3.2.1 mit der relativen Häufigkeit $1/3$ vor, sie konnte nur aus einer nicht gefrorenen $K = 1$ Komponente hervorgegangen sein, da das Abschneiden eines Eingangs des Knotens mit zwei relevanten Eingängen zu einer Komponente führt. Die Komponente aus zwei verbundenen Schleifen (ggf. mit Bäumen) kommt nach Abschn. 3.2.1 mit der relativen Häufigkeit $2/3$ vor, das Abschneiden eines Eingangs des Knotens mit zwei relevanten Eingängen führt zu zwei nicht gefrorenen $K = 1$ Komponenten mit Wahrscheinlichkeit $1/2$. Die mittlere Anzahl der nicht gefrorenen $K = 1$ Komponenten, die durch einen Knoten mit zwei relevanten Eingängen verbunden werden beträgt folglich $1 * (1/3 + 1/2 * 2/3) + 2 * (1/2 * 2/3) = 4/3$.

Die Abschätzung für die mittlere Anzahl der nicht gefrorenen $K = 1$ Komponenten mit relevanten Knoten mit zwei relevanten Eingängen lautet schließlich:

$$\bar{n} = \sum_{m=1}^N p(m)n(m) \approx \sum_{m=1}^N p(m) \ln(2m(e-1)) . \quad (4.57)$$

Da fast alle Knoten mit zwei relevanten Eingängen in einer einzigen komplexen Komponente sitzen, lautet die Abschätzung für die mittlere Anzahl der relevanten Komponenten für große $K = 2$ Netzwerke:

$$C_{rel}^{K=2}(N) \approx \bar{C}(N) - \sum_{m=1}^N p(m)(n(m) - 1) = \bar{C}(N) - \overline{\ln m} - (1 - p(0)) \ln(2\lambda) . \quad (4.58)$$

Die Anwendbarkeit dieser Abschätzung (4.58) wurde durch ihre gute Übereinstimmung mit der in numerischen Simulationen bestimmten Anzahl der relevanten Komponenten belegt, siehe Abb. 4.10. Zum Vergleich der Simulationen mit (4.58) wurden die Ergebnisse für $p(m)$ und (4.52) aus dem Abschn. 4.5.1 benutzt.

Die mittlere Anzahl der komplexen relevanten Komponenten pro Netzwerk kann der Abb. 4.10 entnommen werden als die Differenz zwischen der Anzahl der relevanten Komponenten und der Anzahl der einfachen Schleifen. Für $\beta = 1/8$ und $\gamma = 0, 1/4, 1/2$ ergibt sich für die Anzahl der komplexen Komponenten im Mittel entsprechend 0.6, 0.7 und 0.8. Die Wahrscheinlichkeit dafür, dass es keine Knoten mit zwei relevanten Eingängen in einem Netzwerk gibt, ist für die drei Fälle aus (4.52) entsprechend 0.51, 0.32 und 0.27, sodass die mittlere Anzahl der komplexen

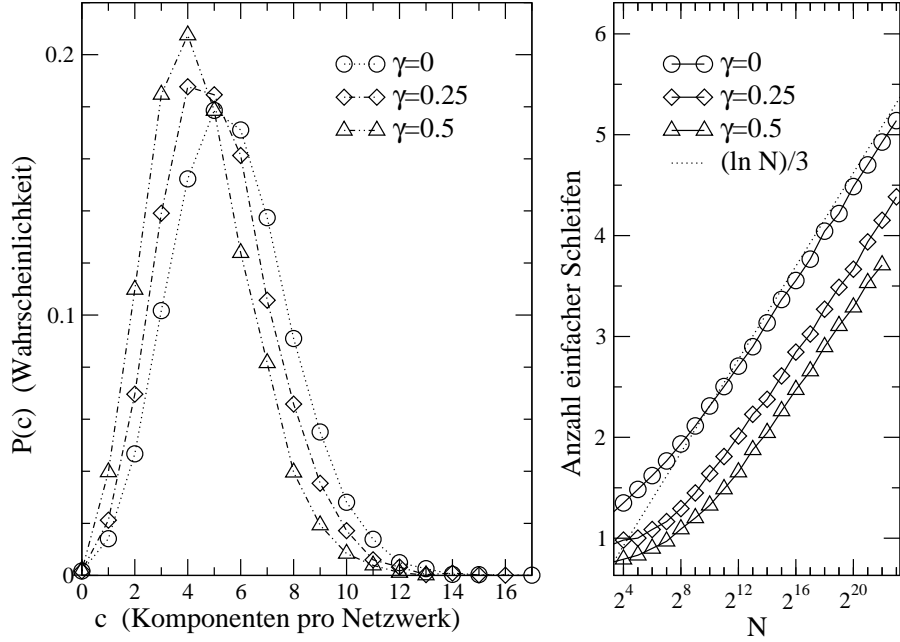


Abb. 4.10.: *Links:* Wahrscheinlichkeitsverteilung für die Anzahl der relevanten Komponenten pro Netzwerk. Die Datenpunkte wurden über 10^4 Netzwerkrealisierungen gemittelt. Benutzte Modellparameter: $\beta = 1/8$; $\gamma = 0, 1/4, 1/2$; $N = 2^{23} \approx 8 \times 10^6$ für $\gamma \neq 1/2$ und $N = 2^{22}$ für $\gamma = 1/2$. Für die mittlere Anzahl der Komponenten ergibt sich in den Simulationen 5.7, 5.1, 4.5 entsprechend für $\gamma = 0, 1/4, 1/2$. Phänomenologisch erhält man mit (4.45) (oder aus der Abb. 4.8) und (4.58) 5.5, 4.5, 3.6. Die Übereinstimmung mit den Simulationen ist zufriedenstellend. Die Halbwertsbreiten der Simulationskurven für dieselben Modellparameter sind 5.3, 5.0, 4.6 im Vergleich zu den mit (4.55) bestimmten Werten 4.43, 4.43, 4.32, wobei die \bar{m} -Abhängigkeit für (4.55) nicht betrachtet wurde. *Rechts:* Anzahl der einfachen Schleifen pro Netzwerk in kritischen RBN mit $K = 2$ für dieselben Modell- und Simulationsparameter wie links. Für die mittlere Anzahl der einfachen Schleifen in den größten betrachteten Netzwerken ergibt sich in den Simulationen, vergleiche (4.53), 5.1, 4.4, 3.7 entsprechend für $\gamma = 0, 1/4, 1/2$. Phänomenologisch erhält man mit (4.58) und (4.52) 5.0, 4.2, 3.3.

relevanten Komponenten in den Netzwerken mit komplexen Komponenten (also nicht in allen) 1.2, 1.0 und 1.1 beträgt. Damit bestätigen die Simulationen die Annahme, dass es meistens nicht mehr als eine komplexe Komponente gibt. Statistische Fluktuationen in den Simulationsergebnissen führen konsistent zu Abweichungen in der ersten Stelle nach dem Komma.

Wir stellen fest, dass die Übereinstimmung der Simulationsergebnisse in Abb. 4.10 mit deren Abschätzung mit (4.58) für kleinere Werte von γ , wo die zur Herleitung von (4.58) benutzten Näherungen plausibler erscheinen, besser ist. Analoge Näherungen, die die kleine Wahrscheinlichkeit dafür berücksichtigen würden, dass die Knoten mit zwei relevanten Eingängen in mehr als eine relevante komplexe Komponente organisiert sein können, würden die Formel (4.57) nicht verändern und zu etwas kleineren Werten für $C_{rel}^{K=2}(N, \beta, \gamma)$ in (4.58) führen. Diese Beobachtung erklärt teilweise die kleine Abweichung der phänomenologischen Abschätzungen von den Simulationen in Abb. 4.10.

4. Kritische RBN mit $K = 2$

Teil II.

Dynamik der RBN mit $K = 2$

Die Erkenntnisse über die im ersten Teil dieser Arbeit studierte Struktur der kritischen booleschen Zufallsnetzwerke mit $K = 2$ werden im zweiten Teil zum Studium der dynamischen Eigenschaften herangezogen. Für einfachere relevante Netzwerkkomponenten als Bestandteile der allgemeinen Netzwerke wurde im ersten Teil die topologische Struktur aufgedeckt. Die Festlegung der Knotenfunktionen in diesen Komponenten spiegelt sich in einer weiteren Struktur wieder, die die Dynamik maßgeblich beeinflusst. Sowohl diese neue Struktur als auch die Dynamik werden im zweiten Teil dieser Arbeit diskutiert. Die wichtigste Erkenntnis besteht dabei darin, dass ein intuitives quantitatives Verständnis der dynamischen Eigenschaften in großen Netzwerken eine direkte Folge der analytischen Ergebnisse über die strukturellen Eigenschaften ist.

Wenn analytische Überlegungen, die man beispielsweise zur detaillierten Betrachtung der Wahrscheinlichkeitsverteilung der Attraktorlängen braucht, intuitiv nur schwer verständliche Aussagen ergeben oder zu kompliziert werden, dann können numerische Simulationen hilfreich sein. Numerische Simulationen werden natürlich auch durchgeführt um analytische Ergebnisse zu bestätigen. Andererseits stoßen bei komplexen Systemen wie den booleschen Netzwerken die numerischen Methoden auf ihre Grenzen, sodass analytische Überlegungen für die Interpretation der Simulationen unabdingbar werden. Analytische Überlegungen und numerische Simulationen werden im zweiten Teil dieser Arbeit sinnvoll kombiniert.

In den Kapiteln 5 und 6 werden die Zahl und Länge der Attraktoren⁵ sowie die entsprechenden Wahrscheinlichkeitsverteilungen besprochen. Im Abschn. 6.1 werden zusätzlich die nicht auf allen Attraktoren gefrorenen Knoten diskutiert. Es wird ausschließlich die synchrone Aktualisierung betrachtet.

Bereits die im Kapitel 5 im Detail beschriebenen einfacheren Netzwerkkomponenten können sehr viele sehr lange Attraktoren haben, was wichtige Implikationen für das Verständnis des RBN - Modells mit $K = 2$ im Kapitel 6 hat. Die einfachen Netzwerkkomponenten zeigen verschiedene komplexe Verhaltensmuster. Sie können daher als geeignet für eine generische Modellierung der genetischen regulatorischen Module angesehen werden, vergleiche dazu im Kapitel 2. Die Beschreibung der einfachen Netzwerkkomponenten erweist sich als kompliziert und trickreich, sodass, wie bei anderen komplexen Systemen, die in einer Reihe von Fällen anwendbaren von den Struktureigenschaften implizierten Untersuchungsmethoden in den anderen Fällen versagen und durch detaillierte Einzelfall - Untersuchungen ergänzt werden müssen.

Im Kapitel 6 werden die dynamischen Eigenschaften der kritischen booleschen Zufallsnetzwerke mit $K = 2$ untersucht. Das wichtigste Ergebnis aus diesem Kapitel besagt, dass sowohl die mittlere Anzahl als auch die mittlere Länge der Attraktoren mit der Netzwerkgröße exponentiell skaliert. Diese Skalierung wurde erst in den

⁵Üblicherweise werden die Attraktoren — sie haben einen Einzugsbereich und sind bei nicht deterministischen Aktualisierungsschemata nicht periodisch — von den periodischen Bahnen bzw. zyklischen Trajektorien bzw. Zyklen im Zustandsraum unterschieden. Der Einfachheit halber weiche ich manchmal von dieser Definition ab und spreche in jedem Fall von den Attraktoren der Dynamik.

letzten Jahren bewiesen, siehe Abschn. 2.4, ihr intuitives Verständnis stammt aus dieser Arbeit.

In numerischen Simulationen werden im Kapitel 6 kritische $K = 1$ Netzwerke (\mathcal{C}_1 Netzwerke), kanalisierende $K = 2$ Netzwerke (\mathcal{C}_2 Netzwerke) und sRBN verglichen. Es wird argumentiert, dass die ähnliche Struktur dieser nicht selten als verschieden angesehenen Modelle zum weitgehend analogen Verhalten führt. Die Interpretation und die Grenzen der Anwendbarkeit der numerischen Simulationen werden unter Ausnutzung des Strukturwissens besprochen.

Die folgenden Kapitel schließen an die Betrachtung im ersten Teil dieser Arbeit an und basieren auf der Arbeit von mir und den Mitverfassern von [Kauf 05a] (Kapitel 5) und [Paul 06] (Kapitel 6) in der Gruppe von Prof. Barbara Drossel an der TU Darmstadt.

5. Einfache Netzwerkkomponenten

Nach einer kurzen Zusammenfassung der bekannten Ergebnisse für die Dynamik der einfachen Schleifen im Abschn. 5.1 werden die Netzwerkkomponenten mit einer Verzweigung¹ im Abschn. 5.2 studiert. Ich zeige, dass die mittlere Anzahl der Attraktoren bereits für diese einfachsten nicht trivialen Netzwerkkomponenten schneller als jedes Potenzgesetz mit der Größe der Komponente wächst. Falls in einer Schleife mit einer Querverbindung dem Verzweigungsknoten mit zwei Eingängen eine reversible Funktion zugeordnet wird, steigt sogar die mittlere Attraktorlänge mit der Anzahl der Knoten exponentiell an. Eine Zusammenfassung weiterer Details zu diesen Ergebnissen findet man am Anfang des Abschn. 5.2.

Es stellt sich heraus, dass das Beispiel der Netzwerkkomponenten mit einer Verzweigung sehr gut dafür geeignet ist, die Grenzen der Anwendbarkeit der numerischen Simulationen im Abschn. 5.2.3 zu diskutieren.

Die Netzwerkkomponenten mit zwei und mehr Verzweigungen können, siehe Abschn. 3.2.1, methodisch analog zum in diesem Kapitel vollständig beschriebenen Fall einer Verzweigung studiert werden. Ich erwarte dabei keine qualitativ neuen Ergebnisse. Während ich mich mit der synchronen Aktualisierung beschäftige, erhält man im Falle der asynchronen stochastischen Aktualisierung qualitativ andere Ergebnisse, die in [Grei 05, Klem 05b] nachgelesen werden können. Dort wird festgestellt, dass die mittlere Anzahl der Attraktoren wie ein Potenzgesetz mit der Komponentengröße skaliert.

Die Schleife mit einer Querverbindung kommt interessanterweise häufig in realen genetischen regulatorischen Systemen vor. Ein möglicher Grund dafür, die unter bestimmten Voraussetzungen erhöhte Zuverlässigkeit der Signalübertragung, wird in [Bran 05] diskutiert. Eine Gegenüberstellung der Ergebnisse aus diesem Kapitel und aus [Bran 05] beziehungsweise aus ähnlichen Arbeiten wäre sicherlich aufschlussreich.

5.1. Einfache Schleifen

Ich fasse hier die (trivialen) dynamischen Eigenschaften der einfachen Schleifen mit N Knoten zusammen um sie zu den Vergleichen im nächsten Abschnitt heranzuziehen.

Wie im Abschn. 3.1 beschrieben, werden gerade und ungerade Schleifen unterschieden. Die dynamischen Eigenschaften lauten [Alda 03b, Flyv 88b]:

¹Eine Verzweigung entspricht zwei Verzweigungsknoten, siehe Kapitel 3

5. Einfache Netzwerkkomponenten

1. Der Anfangszustand wiederholt sich in einer geraden Schleife nach spätestens N Aktualisierungsschritten. Eine ungerade Schleife kehrt zum Anfangszustand spätestens nach $2N$ Aktualisierungsschritten zurück.
2. Folglich gehört jeder Netzwerk - Zustand einer zyklischen Trajektorie im Zustandsraum an und die mittlere Zykluslänge mal die Anzahl der Zyklen ergibt die Mächtigkeit des Zustandsraums 2^N .
3. Gerade Schleifen haben zwei Fixpunkte. Falls keine Knotenfunktion die Negierung ist, dann gibt es die Fixpunkte: alle Knoten auf 0 und alle Knoten auf 1. Ungerade Schleifen haben einen oszillierenden Zyklus der Länge 2 (alternierende 0 und 1 bei einer Negierung).
4. Falls N eine Primzahl ist, beträgt die Anzahl der Zyklen

$$C_N = \begin{cases} 2 + \frac{2^N - 2}{N} & \text{für gerade Schleifen ,} \\ 1 + \frac{2^N - 2}{2N} & \text{für ungerade Schleifen .} \end{cases} \quad (5.1)$$

Ungerade Schleifen mit $N = 2$ bilden eine Ausnahme und besitzen einen Zyklus, der alle vier Zustände einbezieht.

5. Falls N keine Primzahl ist, dann ist, für gerade Schleifen, jeder Teiler von N bzw., für ungerade Schleifen, 2 mal jeder Teiler von N eine realisierbare Zykluslänge. Es gibt vergleichsweise mehr kleinere Zyklen und entsprechend mehr Zyklen als in (5.1) angegeben.

Die Zykluslängen der einfachen Schleifen betragen also größenordnungsmäßig N und die Anzahl der Zyklen — $2^N/N$. Die Anzahl der Zyklen wächst schneller als jede Potenz mit N .

5.2. Komponenten mit Verzweigungen

Die Ergebnisse aus diesem Abschnitt wurden in [Kauf 06, Kauf 05a] veröffentlicht. In den Abschnitten 5.2.1 und 5.2.2 wird im Wesentlichen Folgendes gezeigt.

Es gibt markante Unterschiede im dynamischen Verhalten der Netzwerke mit Verzweigungsknoten mit kanalisierenden und mit reversiblen Funktionen. Insbesondere werden in diesem Abschnitt die mittlere Anzahl und Länge der Zyklen in den Netzwerken mit einer Verzweigung betrachtet. Bei zwei Schleifen mit einer Zwischenverbindung ist die Dynamik im Falle einer reversiblen Funktion an der Verzweigung ähnlich zur Dynamik von zwei unabhängigen Schleifen. Im Falle einer kanalisierenden Funktion hingegen ist die mittlere Anzahl und Länge der Zyklen kleiner und entspricht dem dynamischen Verhalten der einfachen Schleifen. Allgemein ist die Dynamik stark abhängig von der Wahl der Knotenfunktionen.

Je nach Wahl der Parameter findet man charakteristische Zykluslängen. Für zwei Schleifen mit einer Zwischenverbindung findet man eine erhöhte Anzahl der Zyklen

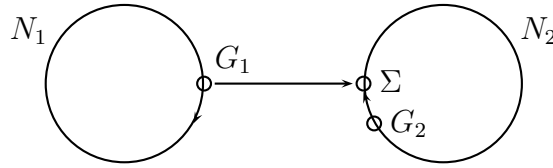


Abb. 5.1.: Zwei Schleifen mit einer Zwischenverbindung.

vor, wenn die Anzahl der Knoten in den Schleifen gleich ist. Eine bestimmte Wahl der kanalisierenden Funktion der Verzweigung führt bei den Schleifen mit einer Querverbindung zur stabilen Dynamik mit einem attraktiven Fixpunkt. Allgemein können die vielfältigen Verhaltensweisen bei einer Auswahl der Netzwerke nach bestimmten Kriterien erzwungen werden.

Die Anzahl der Zyklen steigt bei einer Schleife mit einer Querverbindung exponentiell, jedoch langsamer als bei der entsprechenden einfachen Schleife, mit der Systemgröße an. Die Attraktorlänge beträgt höchstens $2N$ für den Fall kanalisierender Funktionen am Verzweigungsknoten mit zwei Eingängen, während sie für reversible Funktionen exponentiell mit der Schleifengröße ansteigt. Dieses letzte Ergebnis kann direkt auf allgemeinere Netzwerke wie sRBN übertragen werden.

5.2.1. Eine Verzweigung: Zwei Schleifen

In diesem Abschnitt betrachte ich die erste der zwei möglichen Komponenten mit einer Verzweigung aus dem Abschn. 3.2.1, siehe Abb. 5.1. Die Anzahl der Knoten in der ersten und zweiten Schleife bezeichne ich mit N_1 und N_2 . Dem Knoten Σ mit den zwei Eingangsknoten G_1 und G_2 in den beiden Schleifen kann eine der 10 möglichen auf beide Eingänge reagierenden booleschen Funktionen² f_Σ zugeordnet werden. Damit alle Verbindungen relevant sind, müssen den übrigen Knoten einzeln entweder eine Kopierfunktion oder eine Negierung zugeordnet werden.

Symmetrieüberlegungen führen zu Vereinfachungen. Wie für die einfachen Schleifen reicht es für die erste Schleife aus, den Fall einer geraden und einer ungeraden Schleife zu betrachten. Die Komponenten mit N zusätzlichen Knoten zwischen den beiden Schleifen kann man auf die sonst gleiche Komponente mit $N = 0$ abbilden, indem man den Knoten Nummer N , im Uhrzeigersinn ab dem Knoten G_1 gezählt, mit Σ direkt verbindet. Ich beschränke mich auf den Fall der direkten Verbindung zwischen G_1 und Σ . Die zweite Schleife habe außer f_Σ o.B.d.A. nur die Kopierfunktionen. Die Negierungen können nämlich durch eine passende Modifikation von f_Σ mit berücksichtigt werden. Für eine gegebene Funktion f_Σ wird durch Umkehrung aller Knotenzustände oder derjenigen in der ersten Schleife effektiv der Fall einer

²Diese Funktionen sind explizit in der Tabelle 4.1 im Abschn. 2.3.2 aufgeführt

5. Einfache Netzwerkkomponenten

anderen Funktion f_Σ mit erfasst, sodass ich mich von den zehn möglichen auf die drei folgenden Funktionen an der Verzweigung beschränken kann³:

- $f_{11} = G_1 + \bar{G}_2$ (kanalisierend) ,
- $f_{14} = G_1 + G_2$ (kanalisierend) ,
- $f_9 = G_1 \cdot G_2 + \bar{G}_1 \cdot \bar{G}_2$ (reversibel) .

Im Falle der reversiblen Funktion hat jeder Komponentenzustand einen eindeutigen Vorgänger, jeder Komponentenzustand liegt damit auf einem Zyklus (es gibt keine Einzugsbereiche), das Produkt der mittleren Zykluslänge und der Anzahl der Zyklen beträgt $2^{N_1+N_2}$.

Fall 1: Primzahlen N_1 und N_2

Im Folgenden wird angenommen, dass $N_1 \neq N_2$ Primzahlen sind. Die Periode p_1 der Wertefolge am Eingang G_1 von Σ kann N_1 , $2N_1$, 1 und 2 betragen (siehe Abschn. 5.1).

Die zweite Schleife verhält sich wie eine unabhängige einfache Schleife, deren effektive Funktion f'_Σ sich periodisch mit der Periode p_1 ändert: Sie kann mal die Kopierfunktion, mal die Negierung für $f_\Sigma = f_9$ sein; sie kann die Negierung sein oder den gefrorenen Wert 1 annehmen für f_{11} ; und sie kann die Kopierfunktion und 1 für f_{14} sein.

Der Zustand der zweiten Schleife wiederholt sich nach spätestens $2p_1N_2$ Aktualisierungen. Der längste Zyklus hat damit die Periode $4N_1N_2$ und wird erreicht bei einer ungeraden ersten Schleife und $f_\Sigma = f_9$. Im Sonderfall $N_2 = 2$ beträgt die längste Periode $4N_1$.

Für eine kanalisierende Funktion f_Σ folgen viele dynamische Eigenschaften aus der Tatsache, dass für $p_1 > 1$ der Eingang G_1 zumindest alle $2N_1$ Zeitschritte zu 1 wird.

Ich fange mit der Funktion f_{14} an. Einmaliges $G_2 = 1$ führt nach N_2 Aktualisierungen wiederholt zu $G_2 = 1$. Die Kombination $G_2 = 0$ und $G_1 = 0$ führt zu $G_2 = 0$ nach N_2 Zeitschritten. Jedoch wird diese 0 wegen der erwähnten Eigenschaft der ersten Schleife zu 1, spätestens nach $2N_1$ Aktualisierungszyklen der Länge N_2 . Folglich friert nach dieser Zeit die ganze zweite Schleife auf dem Wert 1 ein. Die Zyklen der Komponente haben alle die Länge p_1 ($p_1 > 1$). Für $p_1 = 1$ entstehen Zyklen der Länge N_2 und 1. Für die Anzahl der Zyklen ergibt sich für eine gerade erste Schleife:

Länge	1	N_2	N_1
Anzahl	3	$\frac{2^{N_2}-2}{N_2}$	$\frac{2^{N_1}-2}{N_1}$

Falls die erste Schleife ungerade ist, dann gibt es einen Zyklus der Länge 2 und $(2^{N_1} - 1)/2N_1$ Zyklen der Länge $2N_1$. Der letztere Fall ist ein Beispiel wie alle

³Wieder steht + für logisches ODER, · für logisches UND und \bar{x} für die logische Negation von x .

Knoten in einem Teil des Netzwerks (in der zweiten Schleife) auf allen Zyklen dynamisch einfrieren können.

Die zu unterscheidenden Fälle für $f_\Sigma = f_{11}$ sind:

- Bei $p_1 = 1$ friert die ganze Komponente auf dem Wert 1 ein, sobald 1 der Zustand der ersten Schleife ist.
- Bei $p_1 = 1$ ist die zweite Schleife eine unabhängige ungerade Schleife, falls 0 der Zustand der ersten Schleife ist.
- Bei $p_1 = 2$ befindet sich die ganze Komponente auf einem Zyklus der Länge 2.
- Bei $p_1 = N_1$ ($2N_1$) versklavt die erste Schleife die zweite, das heißt, sie bestimmt den Zustand der zweiten Schleife und löscht alle Information über ihren Anfangszustand aus. Die Zykluslänge beträgt N_1 ($2N_1$) für eine gerade (ungerade) erste Schleife.

Für die Anzahl der Zyklen ergibt sich für eine gerade erste Schleife:

$$\begin{array}{c|c|c|c|c} \text{Länge} & 1 & 2 & 2N_2 & N_1 \\ \hline \text{Anzahl} & 1 & 1 & \frac{2^{N_2}-2}{2N_2} & \frac{2^{N_1}-2}{N_1} \end{array} .$$

Für eine ungerade erste Schleife gibt es einen Zyklus der Länge 2 und $(2^{N_1}-2)/2N_1$ Zyklen der Länge $2N_1$.

Die Anzahl der Zyklen für ungerade Primzahlen N_1 und N_2 und eine kanalisierende Funktion f_Σ ist, zusammengefasst:

$$\begin{aligned} C_{N_1, N_2}^{f_{14}} &= \begin{cases} 3 + \frac{2^{N_1}-2}{N_1} + \frac{2^{N_2}-2}{N_2} & \text{für gerade erste Schleife ,} \\ 1 + \frac{2^{N_1}-2}{2N_1} & \text{für ungerade erste Schleife ,} \end{cases} \\ C_{N_1, N_2}^{f_{11}} &= \begin{cases} 2 + \frac{2^{N_1}-2}{N_1} + \frac{2^{N_2}-2}{2N_2} & \text{für gerade erste Schleife ,} \\ 1 + \frac{2^{N_1}-2}{2N_1} & \text{für ungerade erste Schleife .} \end{cases} \end{aligned} \quad (5.2)$$

Im Sonderfall $N_i = 2$ mit $i \in \{1, 2\}$ müssen die Summanden mit $2N_i$ im Nenner weggelassen werden.

Für große N_1 und N_2 wächst die mittlere Länge der Zyklen asymptotisch linear mit N_{\max} , der größeren der zwei Anzahlen der Schleifenknoten. Die mittlere Anzahl der Zyklen beträgt

$$C_{N_1, N_2}^{f \text{ kanalisierend}} \sim 2^{N_{\max}} / N_{\max} . \quad (5.3)$$

Als Nächstes wird die reversible Funktion f_9 betrachtet, zuerst die einfacheren Fälle. Für eine gerade erste Schleife im Fixpunkt 1 (0) stellt die zweite Schleife eine unabhängige gerade (ungerade) Schleife dar. Die gesamte Komponente hat dann 2 Fixpunkte (einen Zyklus der Länge 2) und $(2^{N_2}-2)/N_2$ Zyklen der Länge N_2 ($(2^{N_2}-2)/2N_2$ Zyklen der Länge $2N_2$). Für eine ungerade erste Schleife auf dem

5. Einfache Netzwerkkomponenten

Zyklus der Länge 2, die gesamte Komponente hat einen Zyklus der Länge 4 und $(2^{N_2} - 2)/2N_2$ Zyklen der Länge $4N_2$.

Wenn die erste Schleife sich auf einem Zyklus der Länge $p_1 = N_1 > 1$ befindet und eine gerade Anzahl der 0en hat, dann wiederholt sich der Wert von G_2 alle N_1N_2 Zeitschritte. Für jeden Zyklus der ersten Schleife der Länge N_1 können in diesem Fall zwei Zyklen der Länge N_1 für die gesamte Komponente folgendermaßen konstruiert werden. Wähle den Anfangswert eines Knotens der zweiten Schleife. Nach einer Aktualisierung wird dadurch der Wert des nächsten im Uhrzeigersinn Knotens in der zweiten Schleife festgelegt usw. Aktualisiere N_1 Mal und notiere den Knotenwert, der dann festgelegt wird, nehme diesen Knotenwert als den Anfangswert des entsprechenden Knotens. Eine N_2 -malige Wiederholung der letzten Anweisung führt zur Festlegung der Anfangswerte aller Knoten der zweiten Schleife. Man kehrt dabei zum ursprünglich gewählten Knoten zurück, der wegen der geraden Anzahl der 0en in der ersten Schleife konsistent auf seinen ursprünglichen Wert festgelegt wird. Somit haben wir einen Zustand eines Zyklus der Länge N_1 konstruiert. Der zweite Zyklus der Länge N_1 ergibt sich analog durch die andere Wahl des ursprünglichen Anfangswerts. Alle anderen Zyklen der Komponente haben die Länge N_1N_2 .

Wenn sich die erste Schleife auf einem Zyklus der Länge $p_1 > 1$ befindet und eine ungerade Anzahl der 0en hat — für ungerade erste Schleifen ist dies immer der Fall — dann wiederholt sich der Wert von G_2 alle $2p_1N_2$ Zeitschritte. Für jeden Zyklus der ersten Schleife der Länge p_1 kann in diesem Fall wie oben beschrieben ein Zyklus der Länge $2p_1$ für die gesamte Komponente konstruiert werden, wobei zwei nacheinander folgende Perioden der ersten Schleife zusammen eine gerade Anzahl von 0en am Knoten G_1 zur Folge haben. Die übrigen Zyklen der Komponente sind $2p_1N_2$ Zustände lang.

Eine Zusammenfassung der Ergebnisse für die Zyklen im Falle einer reversiblen Funktion f_Σ lautet:

Länge	1	2	N_2	$2N_2$	N_1	$2N_1$	N_1N_2	$2N_1N_2$
Anzahl	2	1	$\frac{2^{N_2}-2}{N_2}$	$\frac{2^{N_2}-2}{2N_2}$	$\frac{2^{N_1}-2}{N_1}$	$\frac{2^{N_1}-2}{2N_1}$	$\frac{(2^{N_1}-2)(2^{N_2}-2)}{2N_1N_2}$	$\frac{(2^{N_1}-2)(2^{N_2}-2)}{4N_1N_2}$

für eine gerade erste Schleife und

Länge	4	$4N_2$	$4N_1$	$4N_1N_2$
Anzahl	1	$\frac{2^{N_2}-2}{2N_2}$	$\frac{2^{N_1}-2}{2N_1}$	$\frac{(2^{N_1}-2)(2^{N_2}-2)}{4N_1N_2}$

für eine ungerade erste Schleife.

Im Sonderfall $N_i = 2$ mit $i \in \{1, 2\}$ gilt diese Zusammenstellung nicht. Für $N_1 = 2$ und gerade erste Schleifen findet man keine Zyklen der Längen N_1 und N_1N_2 und die Zyklen der Längen $2N_1$ und $2N_1N_2$ kommen zweimal so oft vor. Für ungerade erste Schleifen muss man die ersten zwei Spalten der obigen Tabelle weglassen und die Anzahl der Zyklen in den übrigen Spalten verdoppeln. Für $N_2 = 2$ und gerade erste Schleifen verschwinden die Spalten Nummer 3, 5 und 7, die Anzahl der Zyklen in den Spalten Nummer 4, 6 und 8 wird verdoppelt. Für ungerade erste

Schleifen verschwinden die Spalten 1 und 3, die Anzahl der Zyklen in den übrigen Spalten wird verdoppelt.

Die mittlere Anzahl der Zyklen für reversible f_Σ divergiert gemäß

$$C_{N_1, N_2}^{f_9} \simeq \begin{cases} \frac{3 \cdot 2^{N_1 + N_2}}{4 N_1 N_2} & \text{für gerade erste Schleife,} \\ \frac{2^{N_1 + N_2}}{4 N_1 N_2} & \text{für ungerade erste Schleife.} \end{cases} \quad (5.4)$$

Die mittlere Länge der Zyklen steigt wie $N_1 N_2$ an. Bis auf den Vorfaktor entspricht dieses Ergebnis dem Fall von zwei unabhängigen Schleifen.

Fall 2: Gleich große $N_1 = N_2 \equiv N$

Das ist der “resonante” Fall, bei dem es für kanalisierende f_Σ und vergleichbare N_1 und N_2 wesentlich mehr Zyklen als im Fall $N_1 \neq N_2$ gibt.

Da jeder Knotenzustand der zweiten Schleife von genau einem Anfangswert in der ersten Schleife (beim Durchgang durch Σ) beeinflusst wird, kann die Komponente in N unabhängige Netzwerke aus zwei Knoten zerlegt werden. Der erste Knoten hat eine regulierende Selbstverbindung, die Negierung bei einer ungeraden ersten Schleife, ansonsten die Kopierfunktion. Der zweite Knoten hängt vom Knotenzustand beider Knoten ab. Diese N Subsysteme bilden eine Folge, deren Elemente eines nach dem anderen aktualisiert werden.

Für eine gerade erste Schleife und $f_\Sigma = f_{14}$ haben die 2 - Knoten - Subsysteme 3 Zyklen der Länge 1. Die gesamte Komponente hat folglich 3 Zyklen der Länge 1 und $\frac{3^N - 3}{N} - \delta_{N,2}$ Zyklen der Länge N .

Für eine ungerade erste Schleife und $f_\Sigma = f_{14}$ hat jedes 2 - Knoten - Subsystem einen Zyklus der Länge 2. Die Komponente hat einen Zyklus der Länge 2 und $\frac{2^N - 2}{2N}$ Zyklen der Länge $2N$. Die erste Schleife versklavt die zweite. Im Sonderfall $N = 2$ gibt es nur einen Zyklus der Länge 4.

Für eine gerade erste Schleife und $f_\Sigma = f_{11}$ hat jedes 2 - Knoten - Subsystem jeweils einen Zyklus der Längen 1 und 2. Die Komponente hat jeweils einen Zyklus der Längen 1 und 2 sowie $\frac{3^N - 3}{2N}$ Zyklen der Länge $2N$. Im Sonderfall $N = 2$ gibt es nur zwei Zyklen der Länge 4.

Für eine ungerade erste Schleife und $f_\Sigma = f_{11}$ versklavt die erste Schleife die zweite. Die Komponente hat einen Zyklus der Länge 2 und $\frac{2^N - 2}{2N}$ Zyklen der Länge $2N$. Im Sonderfall $N = 2$ gibt es nur einen Zyklus der Länge 4.

Für eine gerade erste Schleife und $f_\Sigma = f_9$ hat jedes 2 - Knoten - Subsystem zwei Zyklen der Länge 1 und einen der Länge 2. Die Komponente hat zwei Zyklen der Länge 1, einen Zyklus der Länge 2, $\frac{2^N - 2}{N}$ Zyklen der Länge N (keine für $N = 2$) und $\frac{4^N - 2^N - 2}{2N}$ (3 für $N = 2$) Zyklen der Länge $2N$.

Für eine ungerade erste Schleife und $f_\Sigma = f_9$ hat jedes 2 - Knoten - Subsystem einen Zyklus der Länge 4. Die Komponente hat einen Zyklus der Länge 4 und $\frac{4^N - 4}{4N}$ Zyklen der Länge $4N$. Im Sonderfall $N = 2$ gibt es nur zwei Zyklen der Länge 8.

5. Einfache Netzwerkkomponenten

Asymptotisch divergiert die Anzahl der Zyklen gemäß

$$\begin{aligned} C_{N,N}^{f_9} &\simeq \frac{4^N}{2N} \text{ oder } \frac{4^N}{4N}, \\ C_{N,N}^{f_{14}} &\simeq \frac{3^N}{N} \text{ oder } \frac{2^N}{2N}, \\ C_{N,N}^{f_{11}} &\simeq \frac{3^N}{2N} \text{ oder } \frac{2^N}{2N} \end{aligned} \quad (5.5)$$

entsprechend für eine gerade und eine ungerade erste Schleife. Die mittlere Zykluslänge wächst linear mit N . Meine Simulationen bestätigen die analytischen Ergebnisse, siehe den Anhang B.1.

Fall 3: Allgemeine N_1 und N_2

Angenommen, N_1 und oder N_2 ist keine Primzahl. Dann gibt es vergleichsweise mehr Attraktoren.

Zunächst soll der Fall von N_1 und N_2 , die keine gemeinsamen Teiler besitzen, betrachtet werden. Einfachheitshalber werden für gerade N_2 nur gerade erste Schleifen betrachtet. Die bereits diskutierten Zykluslängen 1, 2, N_1 , N_2 , $2N_1$, $2N_2$, $4N_1$, $4N_2$, N_1N_2 , $2N_1N_2$, $4N_1N_2$ kommen weiterhin vor. Man erhält jedoch auch andere Zykluslängen, indem man N_1 und oder N_2 in dieser Liste durch einen ihrer Teiler ersetzt. Die Anzahl der Zyklen mit Längen aus der ursprünglichen Liste wird entsprechend kleiner.

Im Hauptteil dieses Abschnitts beschäftige ich mich mit dem interessanteren Fall, wo der größte gemeinsame Teiler g von N_2 und der Länge des Zyklus der ersten Schleife P_1 größer 1 ist. Es folgt $g > 1$, wenn N_1 und N_2 einen gemeinsamen Teiler ungleich N_1 haben. (Der Fall $N_1 = N_2$ wurde im vorigen Abschnitt untersucht.)

Das kleinste gemeinsame Vielfache von P_1 und N_2 ist P_1N_2/g . Für ein gegebenes P_1 beträgt die größte mögliche Länge des Zyklus der gesamten Komponente $2P_1N_2/g$. Sie wird für reversible f_Σ erreicht. Die kleinste mögliche Länge beträgt P_1 . Die Folge der G_1 - Knotenwerte während einer Periode der ersten Schleife und die Gesamtheit der Knotenwerte der zweiten Schleife zerfallen in g unabhängige Subsysteme, bestehend jeweils aus einer Sequenz der P_1/g der G_1 - Knotenwerte und der entsprechenden von ihnen beeinflussten N_2/g Knotenwerte der zweiten Schleife. Aktualisierungen der gesamten Komponente sind demnach äquivalent dazu, dass die Subsysteme gemäß ihrer natürlichen Reihenfolge nacheinander aktualisiert werden. Eine Aktualisierung bewirkt mit anderen Worten, dass zuerst ein ausgezeichnetes Subsystem aktualisiert wird und danach die Sequenz der Subsysteme zyklisch um eine Stelle verschoben wird, sodass das nächste Subsystem ausgezeichnet wird. Ein Subsystem S , dass die zyklisch angeordneten P_1/g Knotenwerte der ersten Schleife beinhaltet, kann man sich anschaulich als eine gerade Schleife \mathcal{S}_S aus P_1/g Knoten mit den obigen Knotenwerten vorstellen, wobei die beeinflussten Knotenwerte der zweiten Schleife als die Werte der in eine weitere Schleife angeordneten Knoten vorgestellt werden können.

Bevor ich auf das Zusammenspiel der Subsysteme zu sprechen komme, folgen hier einige Bemerkungen. Für eine ungerade erste Schleife und eine gerade Zahl P_1/g sind die Knotenwerte einer Hälfte von \mathcal{S}_S einfach gleich der Negierung der Knotenwerte der anderen Hälfte. Für eine ungerade erste Schleife und P_1/g ungerade sind die Subsysteme paarweise “komplementär”: Zu jedem Subsystem S_1 mit einer ungeraden Anzahl von 0en in \mathcal{S}_{S_1} existiert ein anderes Subsystem S_2 mit einer geraden Anzahl von 0en in \mathcal{S}_{S_2} .

Die Länge und die Anzahl der Zyklen eines Subsystems S können, wie in den vorangegangenen Abschnitten erklärt, berechnet werden. Ich möchte hier einige Regeln beschreiben, die bei gegebenen Längen der Zyklen p_1, p_2, \dots der einzelnen Subsysteme S_1, S_2, \dots zur Bestimmung der Zykluslängen der gesamten Komponente nützlich sind. Falls alle Subsysteme unterschiedliche Zyklen durchlaufen, beträgt die Länge des Zyklus der gesamten Komponente $T = \text{KGV}(p_1g, p_2g, \dots)$. KGV steht für das kleinste gemeinsame Vielfache. Andernfalls existieren kürzere Zyklen. Beispielsweise können die Phasenverschiebungen zwischen den Subsystemen auf dem gleichen Zyklus $p_1 = p_2 = \dots = p$ so gewählt werden, dass kürzere als pg Zykluslängen der gesamten Komponente vorkommen. Diese Längen können ein Teiler von pg sein, der ein Vielfaches von p , aber kein Vielfaches von g ist.

Die Diskussion der Anzahl der Zyklen beginne ich mit dem Fall $f_\Sigma = f_9$. Bei geraden ersten Schleifen reicht es für große N_1 und N_2 aus, $P_1 = N_1$ zu betrachten. Damit ist jedes Subsystem S mit einer Wahrscheinlichkeit nahe 0.5 auf einem Zyklus der Länge $N_1/g \times N_2/g$. Es ist auf einem Zyklus der Länge $N_1/g \times 2N_2/g$ auch mit einer Wahrscheinlichkeit nahe 0.5. Die Subsysteme befinden sich meistens auf verschiedenen Zyklen. Die Zykluslänge der gesamten Komponente N_1N_2/g kommt dann mit der Wahrscheinlichkeit 0.5^g vor, die Länge $2N_1N_2/g$ kommt mit der Wahrscheinlichkeit $(1 - 0.5^g)$ vor. Die Zyklen der Länge N_1N_2/g können vernachlässigt werden, da ihr Vorkommen $0.5^g/(1 - 0.5^g)$ -mal unwahrscheinlicher im Vergleich zu den Zyklen der Länge $2N_1N_2/g$ ist. Auch die Beiträge der Zyklen der Länge $2N_2, N_2, 2N_1, N_1$ zur Gesamtzahl der Zyklen werden vernachlässigt.

Bei ungeraden ersten Schleifen kann man sich auf den Fall $P_1 = 2N_1$ beschränken. Jedes Subsystem S bzw. jedes Paar der komplementären Subsysteme befindet sich mit einer Wahrscheinlichkeit nahe 1 auf einem Zyklus der Länge $2 \times 2N_1/g \times N_2/g$, so dass die Zykluslänge der gesamten Komponente $4N_1N_2/g$ ist. Bei der Abschätzung der Anzahl der Zyklen entsprechen die gewichtigeren Beiträge, die ich vernachlässige, den Zyklen der Länge $4N_1$ ($P_1 = 2N_1$) und $4N_2/g$ ($P_1 = 2$). Die mit der Gleichung (5.4) zu vergleichende Abschätzung der mittleren Anzahl der Zyklen der Komponente lautet schließlich

$$C_{N_1, N_2}^{f_9} \simeq \begin{cases} \frac{g}{2} \frac{2^{N_1+N_2}}{N_1N_2} & \text{für gerade erste Schleife,} \\ \frac{g}{4} \frac{2^{N_1+N_2}}{N_1N_2} & \text{für ungerade erste Schleife.} \end{cases} \quad (5.6)$$

Für kanalisierende Funktionen f_Σ erhält man für die gerade und die ungerade erste Schleife qualitativ verschiedene Ergebnisse. Die ungerade erste Schleife versklavt für ungerade N_2 die zweite Schleife. Die Ergebnisse sind von N_2 unabhängig

5. Einfache Netzwerkkomponenten

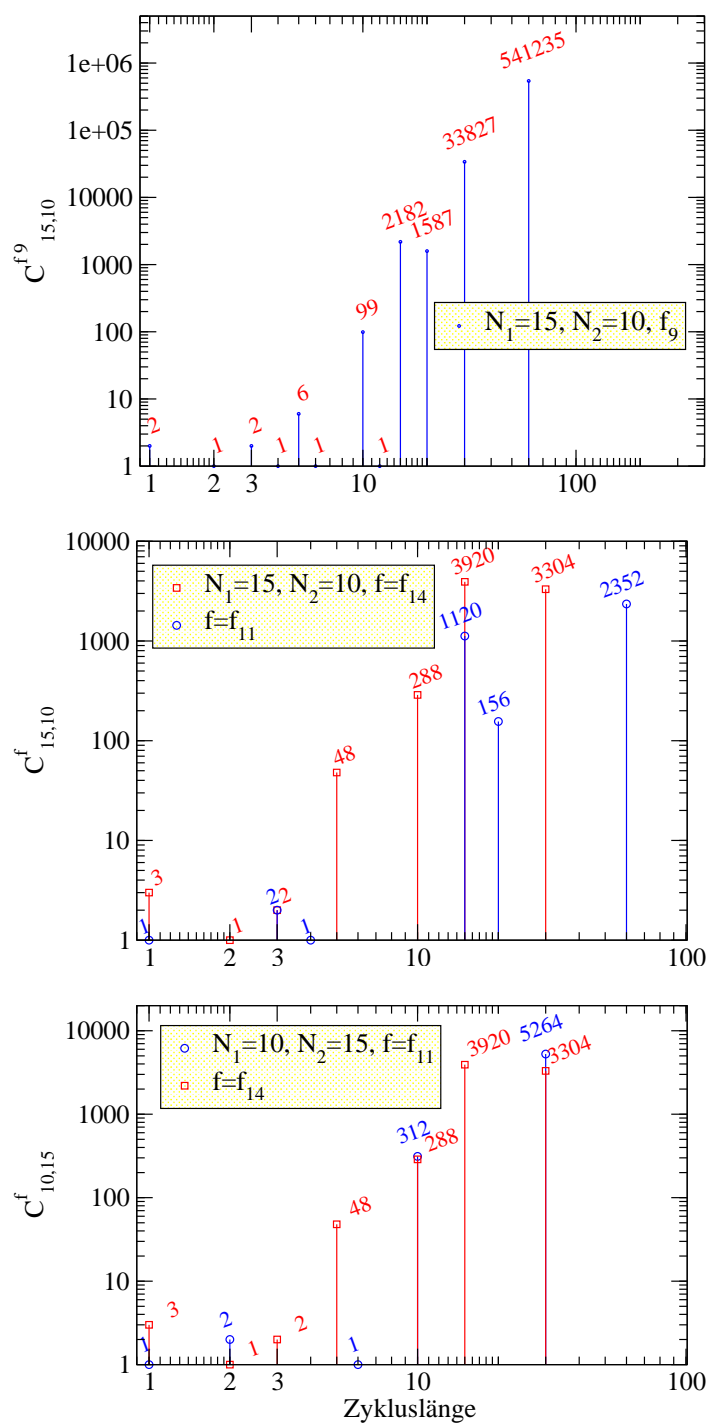


Abb. 5.2.: Simulationsergebnisse für die Abhängigkeit der Anzahl der Zyklen von ihrer Länge für zwei Schleifen mit einer Zwischenverbindung bestätigen analytische Überlegungen. Hier sind drei Beispiele mit einer geraden ersten Schleife dargestellt, $g = 5$.

und ergeben nichts Neues im Vergleich zu den vorangegangenen Unterabschnitten. Die Mehrheit der Zyklen der Komponente haben die Länge $2N_1$, ihre Anzahl ist größenordnungsmäßig $2^{N_1}/2N_1$, vgl. Gleichung (5.2).

Im Fall einer ungeraden ersten Schleife und für gerade N_2 haben P_1 und N_2 den gemeinsamen Teiler 2, sodass zwei (im oben beschriebenen Sinne) komplementäre unabhängige Subsysteme entstehen, die beide nunmehr einer geraden ersten Schleife mit N_1 Knoten entsprechen. Die Anzahl der Zyklen der gesamten Komponente kann bei bekannten analogen Ergebnissen für die beiden Subsysteme aus ihnen verquickt werden. Somit reduziert sich das Problem auf den gleich zu behandelnden Fall einer geraden ersten Schleife. (Die beiden Subsysteme zerfallen bei $g \neq 2$ in weitere unabhängige Subsysteme.) Als Beispiel der angesprochenen Verquickung betrachte ich $P_1 = 2$ und ein gerades N_2 . Im ersten der unabhängigen Subsysteme der Knotenwerte von G_1 und G_2 gilt immer $G_1 = 1$, im zweiten gilt immer $G_1 = 0$. In diesem Sonderfall, wo G_1 in den Subsystemen konstant ist, erhält man größenordnungsmäßig $2^{N_2/2}/(N_2/2)$ Zyklen der gesamten Komponente der Länge N_2 für $f_\Sigma = f_{14}$ und $2^{N_2/2}/N_2$ Zyklen der Länge $2N_2$ für $f_\Sigma = f_{11}$. Allgemein kann die Verquickung langwierig sein, siehe dazu den Text weiter unten zu geraden ersten Schleifen und die Details im Anhang B.2. Im speziellen Grenzfall $N_1 \gg N_2$ ($N_2 \gg N_1$) ist die Größenordnung der mittleren Anzahl der Attraktoren $g^2 2^{N_1}/N_1$ ($g^2 2^{N_2}/N_2$) vergleichbar mit derjenigen im Falle der Primzahlen N_1 und N_2 .

Bei einer geraden ersten Schleife tauchen neue im Vergleich zum Fall der Primzahlen N_1 und N_2 Zykluslängen auf. Es ergibt sich außerdem ein qualitativ anderes Ergebnis für die mittlere Anzahl der Attraktoren. Insbesondere treten Zyklen der Längen $N_1 N_2 / g$ und $2N_1 N_2 / g$ auf, weil einige der unabhängigen Subsysteme Zyklen der Länge N_1 / g und andere — Zyklen der Länge N_2 / g haben können.

Zuerst sei $f_\Sigma = f_{14}$. Für große N_1 und N_2 ist jedes Subsystem fast sicher in einem der ungefähr $2^{N_1/g}$ Zustände, die zu den Zyklen der Länge N_1 / g gehören, oder in einem der ungefähr $2^{N_2/g}$ Zustände der Zyklen der Länge N_2 / g . Für große Schleifen ergibt sich die Anzahl der Zyklen der Komponente zu

$$C_{N_1, N_2}^{f_{14}} \simeq \frac{g}{N_1 N_2} (2^{N_1/g} + 2^{N_2/g})^g. \quad (5.7)$$

Eine detailliertere Betrachtung führt auf den Ausdruck

$$C_{N_1, N_2}^{f_{14}} \simeq \frac{g (2^{N_1/g} + 2^{N_2/g} - 1)^g}{N_1 N_2} + \frac{(2^{N_1/g} + 1)^g}{N_1} + \frac{(2^{N_2/g} + 1)^g}{N_2}, \quad (5.8)$$

wobei nur die dominanten Zykluslängen $N_1 N_2 / g$, N_1 und N_2 berücksichtigt wurden, siehe den Anhang B.2 für mehr Details.

Im Fall $f_\Sigma = f_{11}$ beträgt die größte Zykluslänge $N_1 N_2 / g$ für gerade N_1 / g und $2N_1 N_2 / g$ für ungerade N_1 / g . Die Details findet man im Anhang B.2. Für die Anzahl

5. Einfache Netzwerkkomponenten

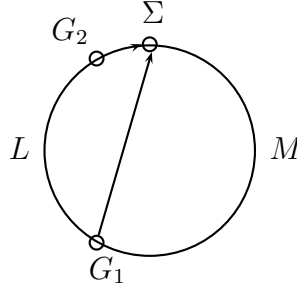


Abb. 5.3.: Schleife mit einer Querverbindung.

der Zyklen ergibt sich

$$C_{N_1, N_2}^{f_{11}} \simeq \begin{cases} \frac{g(2^{N_1/g} + 2^{N_2/g} - 1)^g}{\frac{N_1 N_2}{(2^{N_2/g} + 1)^g} + \frac{2N_2}} + \frac{(2^{N_1/g} + 1)^g}{N_1} & \text{für gerade } N_1/g, \\ \frac{g(2^{N_1/g} + 2^{N_2/g} - 1)^g}{\frac{2N_1 N_2}{(2^{N_2/g} + 1)^g} + \frac{2N_1}} + \frac{(2^{N_1/g} + 1)^g}{2N_1} & \text{für ungerade } N_1/g. \end{cases} \quad (5.9)$$

Zur Veranschaulichung der präsentierten Ergebnisse sind in Abb. 5.2 drei Beispielsimulationen zur Bestimmung der Zyklen der zwei Schleifen mit einer Zwischenverbindung und $g = 5$ dargestellt. Verglichen mit dem Fall von zwei unabhängigen Schleifen ist die größte mögliche Zykluslänge vier Mal so groß wie $N_1 N_2 / g$. Für eine kanalisierende Funktion f_Σ sind die Zyklen vergleichsweise kürzer, dafür aber zahlreicher. Die Komponente mit der Zwischenverbindung hat charakteristische dominante Zykluslängen. Die Anzahl der Zyklen steigt schneller als jede Potenz mit N_1 und N_2 an, die mittlere Zykluslänge wächst linear mit den Größen der Schleifen.

5.2.2. Eine Verzweigung: Querverbindung

Die zweite der zwei nach Abschn. 3.2.1 möglichen Komponenten mit einer Verzweigung wird in diesem Abschnitt besprochen. Die Schleife mit einer Querverbindung ist in Abb. 5.3 dargestellt. Sie besteht aus den Verzweigungsknoten G_1 und Σ sowie aus den weiteren $N - 2 = L + M$ Knoten. Der Knoten Σ erhält zwei eingehende Verbindungen von den Knoten G_1 und G_2 . Allen Knoten wird eine Nummer zugewiesen: Gegen den Uhrzeigersinn, $x = 1$ für G_2 , $x = L + 1$ für G_1 und $x = N \bmod N \equiv 0$ für Σ .

Ohne Beschränkung der Allgemeinheit wird angenommen, dass die Knotenfunktion aller Knoten außer Σ die Kopierfunktion ist, vgl. Abschn. 5.2.1. Des Weiteren kann ich mich von den 10 möglichen auf beide Eingänge reagierenden booleschen Funktionen auf folgende f_Σ beschränken:

- $f_{14} = G_1 + G_2$ (kanalisierend) ,

- $f_4 = \bar{G}_1 \cdot G_2$ (kanalisierend) ,
- $f_1 = \bar{G}_1 \cdot \bar{G}_2$ (kanalisierend) ,
- $f_{11} = G_1 + \bar{G}_2$ (kanalisierend) ,
- $f_9 = G_1 \cdot G_2 + \bar{G}_1 \cdot \bar{G}_2$ (reversibel) .

Um die anderen fünf Funktionen zu berücksichtigen muss man lediglich alle Knotenwerte invertieren. Die Komponente mit $n < L$ Knoten auf der geradlinigen Verbindung zwischen G_1 und Σ kann auf die gleiche Komponente mit $n = 0$ durch direkte Verbindung des Knotens Nummer $L + 1 - n$ mit Σ bijektiv abgebildet werden. Aus diesem Grund wird immer $n = 0$ angenommen.

Falls der größte gemeinsame Teiler g von N und L größer 1 ist, zerfällt die Komponente in g dynamisch unabhängige Subsysteme mit N/g Knoten. Die Zyklen der gesamten Komponente können dann aus den Zyklen der Subsysteme verquickt werden, so wie im Abschn. 5.2.1 für zwei verbundene Schleifen beschrieben. Wenn nicht anders vermerkt, wird bei analytischen Überlegungen in diesem Abschnitt $g = 1$ angenommen, was für die entsprechenden Aussagen keine wesentliche Rolle spielt.

Im Folgenden betrachte ich nacheinander die vier obigen kanalisierenden Funktionen und die reversible Funktion f_9 mit Hilfe von analytischen Überlegungen und Computersimulationen.

Fall 1: $f_\Sigma = f_{14}$

In diesem einfachsten Fall wird der Ausgang von Σ nur dann 0 sein, wenn beide Eingänge 0 sind. Für einen beliebigen Komponenten - Anfangszustand kann die Anzahl der 0en im Laufe der Dynamik nicht kleiner werden. Es gibt zwei Fixpunkte, alle Knoten auf 0 und alle auf 1. Jede 0 braucht eine andere 0, L Schritte zurück entlang der Schleifenverbindungen um zu überleben. Nicht triviale Zyklen bekommt man nur dann, wenn $g > 1$ ist. Die g Subsysteme können unabhängig auf dem Wert 0 oder 1 eingefroren werden. Die Zyklen der Länge g oder der Länge eines der Teiler von g nehmen $2^g - 2$ Zustände im Zustandsraum ein. Die über verschiedene L und ein kleines Werteintervall von N gemittelte Anzahl der Zyklen wächst mindestens wie $2^{N/2}/N^2$ mit N , da $g = N/2$ für gerade N und $L = N/2$ gilt.

Fall 2: $f_\Sigma = f_4$

Die kanalisierende Funktion f_4 ergibt 0, wenn der erste Eingang G_1 1 ist, sonst kopiert sie den Wert des zweiten Eingangs. Für einen beliebigen Anfangszustand kommt jede 0 nach N Aktualisierungsschritten (nach einer Rotation) zurück an ihren ursprünglichen Platz. Auf einem Zyklus muss daher auf jede $G_2 = 1$ eine (L Knoten zurück liegende) $G_1 = 0$ folgen, sonst verschwindet diese 1 nach Durchgang durch Σ . Ich betrachte nun die Sequenz der sich alle L Zeitschritte ergebenden Zustände von G_2 auf einem Zyklus. (Bei $g > 1$ gibt es g unabhängige Sequenzen der

5. Einfache Netzwerkkomponenten

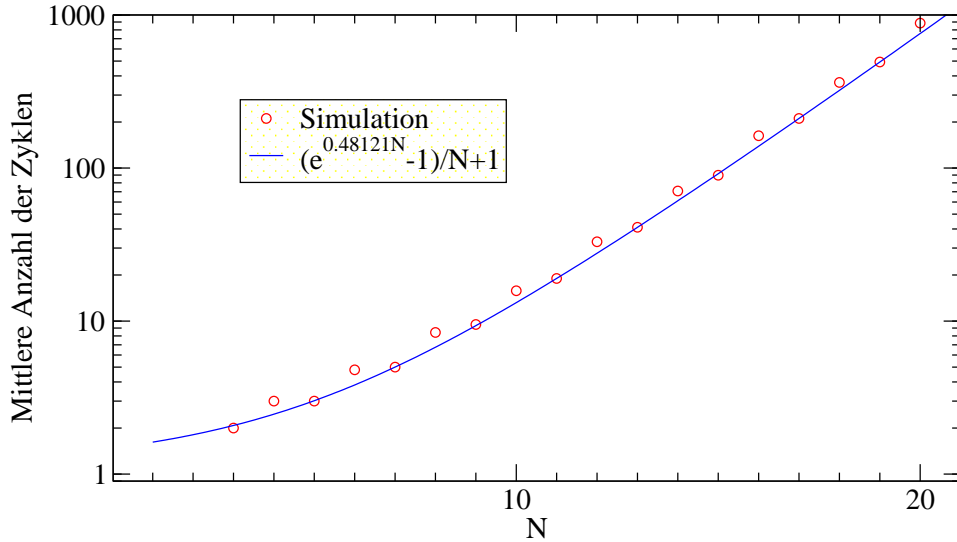


Abb. 5.4.: Mittlere Anzahl der Zyklen der Schleife mit einer Querverbindung im Falle $f_\Sigma = f_4$, gemittelt über alle möglichen L -Werte.

Länge N/g .) Für die Anzahl ϕ_N der verschiedenen möglichen Sequenzen der Länge N , in denen jede 1 von einer 0 gefolgt wird, erhält man die rekursive Gleichung

$$\phi_N = \phi_{N-1} + \phi_{N-2},$$

wobei jede Sequenz der Länge N aus einer Sequenz der Länge $N-1$ durch Einfügen einer 0 nach der ersten 1 (bzw. am Ende, falls keine 1 in der Sequenz vorhanden ist) oder aus einer Sequenz der Länge $N-2$ durch Einfügen von 01 nach der ersten 1 (bzw. 00 am Ende, falls keine 1 in der Sequenz vorhanden ist) entsteht. Die Anfangsbedingung ist $\phi_1 = 1$ und $\phi_2 = 3$. Die exakte Lösung der Rekursionsgleichung lautet $\phi_N = \tau^N + (-1/\tau)^N$ mit $\tau = (1 + \sqrt{5})/2$. Eine gute Näherung für große N ist $\phi_N = \tau^N$. Für die Anzahl der Zyklen, die bis auf den trivialen Beitrag des Fixpunkts gleich der Anzahl der Komponentenzustände auf den Zyklen der Länge N geteilt durch N ist, erhält man bei $g = 1$ den Ausdruck

$$C_N^{f_4} \simeq \frac{e^{0.48121N} - 1}{N} + 1. \quad (5.10)$$

Bei $g > 1$ ist diese Anzahl etwas höher. Das Ergebnis wird durch eine numerische Simulation mit Mittelung über verschiedene L in Abb. 5.4 bestätigt.

Fall 3: $f_\Sigma = f_1$

Die kanalisierende Funktion f_1 produziert eine 1 genau dann, wenn ihre beiden Eingänge 0 sind. Folglich ist der Ausgang 0, sobald einer der Eingänge 1 ist. Die Anzahl der Zyklen steigt im Fall $f_\Sigma = f_1$ wieder exponentiell mit der Komponentengröße an, der Beweis ist hier etwas langwieriger.

Was passiert mit einem Knotenwert 1 auf einem Zyklus im Laufe der Dynamik? Diese 1 springt vom Ort x zum Ort $x - 1$ in einem Zeitschritt. Beim Erreichen des Knotens G_1 führt sie zu $\Sigma = 0$. Beim Erreichen von G_2 produziert sie eine andere 0, genau L Plätze hinter der ersten. Diese zwei 0en produzieren wiederum eine 1, wenn sie die Knoten G_1 und G_2 erreichen, sodass eine 1 nach $2N - L$ Schritten an ihren ursprünglichen Platz zurückkehrt. Jedes Paar von 0en im Abstand von L Knoten voneinander besitzt dieselbe Eigenschaft.

Um zu beweisen, dass die Zykluslänge $2N - L$ oder einer der Teiler von $2N - L$ ist, muss man zeigen, dass jede 0 auf dem Zyklus zu einem Paar von 0en im Abstand von L gehört. Angenommen, am Platz Σ erscheint im nächsten Zeitschritt eine 1. Falls L Zeitschritte später eine 0 sie am Platz Σ ersetzt, muss gleichzeitig $G_1 = 1$ gelten. Nach noch einmal L Zeitschritten folgt dann $\Sigma = 0$, L Schritte hinter der ersten 0, die damit zum Paar gehört.

Für $L = 1$ ist die Anzahl der Komponentenzustände auf den Zyklen gleich der Anzahl der Sequenzen der Länge N mit der durch den Σ -Durchgang bestimmten Randbedingung, in denen die 0en immer in Paaren vorkommen. Die Anzahl dieser Sequenzen ϕ_N erfüllt die Rekursionsgleichung

$$\phi_N = 2\phi_{N-1} - \phi_{N-2} + \phi_{N-3}, \quad N \geq 4. \quad (5.11)$$

Diese Relation ergibt sich wie folgt. Eine Sequenz der Länge N entsteht, wenn eine 1 oder eine 0 nach der ersten 1 in einer Sequenz der Länge $N - 1$ eingefügt wird. Daher kommt der Summand $2\phi_{N-1}$. Wenn auf die erste 1 eine zweite folgt, dann ist es verboten, eine 0 einzufügen. Die Anzahl dieser verbotenen Sequenzen ist ϕ_{N-2} , da sie durch Einfügen einer 1 nach der ersten 1 in einer Sequenz der Länge $N - 2$ entstehen. Der Term ϕ_{N-2} ist also abzuziehen. Um die Sequenzen zu bauen, in denen 001 auf die erste 1 folgt, kann man die drei Bits nach der ersten 1 in einer Sequenz der Länge $N - 3$ einfügen. Daher die Addition des dritten Terms in (5.11). Die Sequenzen, die aus lauter Nullen bestehen oder eine einzige 1 am Ende haben, ergeben sich durch Zusatz einer 0 oder einer 1 zur Sequenz aus lauter Nullen der Länge $N - 1$. Unter Berücksichtigung der Randbedingung wird die Rekursion mit $\phi_1 = 0$, $\phi_2 = 3$, $\phi_3 = 5$ initialisiert.

Die exakte Lösung von (5.11) mit der Randbedingung lautet $\phi_N = \tau(\tau - 1) \times \frac{\cos(\phi N - \phi/2)}{\cos(\phi/2)} \times (\tau^{N-1} - (\tau - 1)^{N-1})$. τ und $(\tau - 1) \exp(\pm i\phi)$ sind die Lösungen der Gleichung $x^3 - 2x^2 + x - 1 = 0$. Eine gute Näherung für große N lautet: $\phi_N \approx (\tau - 1)\tau^N \approx 0.75488 \exp(0.56240N)$. Die Gesamtanzahl der Zyklen kann für $L = 1$ als $\phi_N/(2N - 1)$ abgeschätzt werden.

Bei $L > 1$ wird beim Ablesen der periodischen Sequenz der Werte der Knoten, die jeweils im Abstand von L Knoten voneinander liegen, der Σ -Knoten L -mal überschritten, sodass die Bestimmung der "Randbedingungen" umständlich ist. In Simulationen beobachte ich, dass für große L der Faktor im Exponenten kleiner ist. Die numerische Auswertung der über die möglichen L gemittelten Anzahl der Zyklen wird in Abb. 5.5 dargestellt. Das asymptotische Verhalten ist für die ausgewerteten kleinen N noch nicht zu sehen.

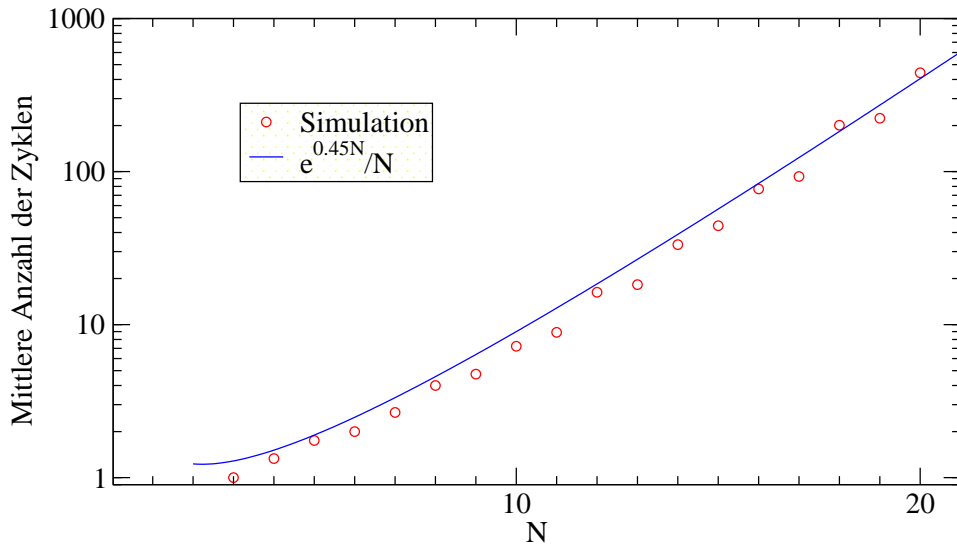


Abb. 5.5.: Mittlere Anzahl der Zyklen der Schleife mit einer Querverbindung im Falle $f_{\Sigma} = f_1$, gemittelt über alle möglichen L -Werte. Das asymptotische Verhalten wird erst bei größeren N erreicht.

Fall 4: $f_{\Sigma} = f_{11}$

Die letzte zu diskutierende kanalisierende Funktion f_{11} produziert eine 1, wenn der erste Eingang $G_1 = 1$, sonst gibt sie den invertierten Wert des zweiten Eingangs zurück. Mit anderen Worten ist der Ausgang genau dann 0, wenn $G_1 = 0$ und $G_2 = 1$. Die Schleife mit $f_{\Sigma} = f_{11}$ hat einen Fixpunkt für alle Knotenwerte gleich 1.

Wir verfolgen das Schicksal einer 1 auf einem Zyklus im Laufe der Dynamik. Diese 1 springt vom Ort x zum Ort $x - 1$ in einem Zeitschritt. Beim Erreichen des Knotens G_1 führt sie zu $\Sigma = 1$. Folglich kommt die 1 nach $N - L$ Schritten an ihren ursprünglichen Platz zurück. Analog dazu kommt jede 0 nach $N - L$ Schritten an ihren ursprünglichen Platz zurück, wenn eine 1 genau L Zeitschritte zurück an diesem Platz war.

Um zu zeigen, dass die Zykluslänge tatsächlich $N - L$ oder ein Teiler von $N - L$ ist, muss bewiesen werden, dass auf einem Zyklus jede 0 in L Schritten von einer 1 gefolgt wird. Angenommen, $\Sigma = 0$. Diese 0 kann dynamisch entstanden sein, wenn der Wert des Knotens mit der Nummer L gleichzeitig 0 ist. L Zeitschritte später wird daher $\Sigma = 1$ folgen.

Zur Abschätzung der Anzahl der Zyklen wird die Sequenz aus den Knotenwerten von G_1 zu N Zeitpunkten im Abstand von L Aktualisierungen auseinander betrachtet, also für N solche Zeitintervalle in Folge. Die Anzahl der möglichen Sequenzen ist gleich der Anzahl der Komponentenzustände auf den Zyklen. Im einfachsten Fall

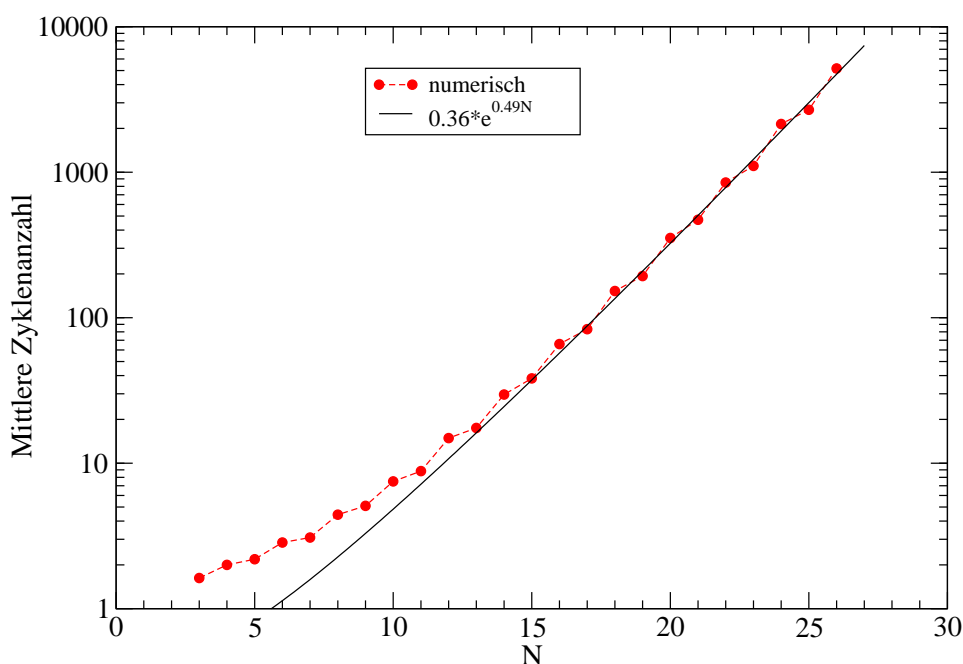


Abb. 5.6.: Mittlere Anzahl der Zyklen der Schleife mit einer Querverbindung für kanalisierende Funktionen f_Σ , gemittelt auch über alle möglichen L -Werte. Die analytischen Überlegungen im Haupttext legen nahe, dass das asymptotische Verhalten erst bei größeren N erreicht wird.

$L = 1$ ergibt sich für diese Sequenzen eine Randbedingung wegen einem Durchgang durch Σ . Nach Berücksichtigung der Randbedingung dürfen in diesen Sequenzen außerdem die Nullen nicht in Paaren auftreten. Die Anzahl der Sequenzen erfüllt die Rekursionsgleichung $\phi_N = \phi_{N-1} + \phi_{N-2}$, da eine Sequenz der Länge N genau dann entsteht, wenn entweder eine 1 nach der ersten 0 in einer Sequenz der Länge $N - 1$ oder die Bitfolge 10 nach der ersten 0 in einer Sequenz der Länge $N - 2$ eingefügt wird. Man kann zeigen, dass diese Rekursionsgleichung auch für $L \ll N$ gilt, nur der Vorfaktor in der Lösung ändert sich. Die Rekursionsgleichung ist identisch mit derjenigen im Fall $f_\Sigma = f_4$, sodass die Gesamtanzahl der Zyklen bei kleinen L wieder wie $e^{0.48121N}/N$ divergiert.

Die Ergebnisse für alle vier untersuchten kanalisierenden Funktionen deuten darauf hin, dass die mittlere Anzahl der Zyklen der Schleife mit einer Querverbindung, gemittelt über alle L und alle kanalisierenden Funktionen, mindestens wie $e^{0.5624N}/N^2$ ansteigt, denn größenordnungsmäßig der Anteil gleich $1/N$ aller Komponenten größenordnungsmäßig $e^{0.5624N}/N$ Zyklen besitzt. Dieses Verhalten wird bei den in meinen Computersimulationen zugänglichen Werten von N jedoch noch nicht erreicht, siehe Abb. 5.6.

5. Einfache Netzwerkkomponenten

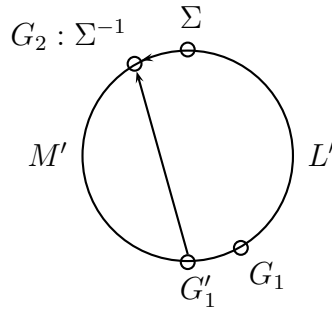


Abb. 5.7.: Die Komponente mit der umgekehrten Zeitrichtung im Vergleich zu derjenigen in Abb. 5.3 mit einer reversiblen booleschen Funktion f_Σ . Die Knoten G_1 und G_1' sowie die Knoten G_2 und Σ sind direkt verbunden. Die Knotenfunktion von G_2 macht die Einwirkung der Funktion f_Σ rückgängig.

Fall 5: $f_\Sigma = f_9$

Für eine reversible boolesche Funktion des Knotens Σ ist die Dynamik reversibel. Alle Komponentenzustände liegen auf Zyklen. Da durch die Zeitumkehr eine Komponente mit $L \leq M + 2$ auf eine mit $L > M + 2$ abgebildet werden kann, ist es ausreichend, den Fall $L \leq M + 2$ bzw.

$$1 \leq L \leq [N/2] \quad (5.12)$$

zu betrachten. Die Abb. 5.7 zeigt die entsprechende Komponente mit der umgekehrten Zeitrichtung, wobei

$$\begin{aligned} L' &= M + 2 = N - L, \\ M' &= L - 2. \end{aligned} \quad (5.13)$$

Zu beachten ist, dass $M' = -1$ ist, wenn die Querverbindung eine Selbstverbindung ist.

Analog zum Fall der kanalisierenden Funktionen f_Σ zerfallen die Knoten in g unabhängige Subsysteme mit N/g Knoten, falls N und L kommensurabel sind und g ihr größter gemeinsamer Teiler ist. Im Gegensatz zum Fall der kanalisierenden Funktionen f_Σ liegt für reversible Funktionen jeder Komponentenzustand auf einem Zyklus. Das führt dazu, dass sehr lange Zyklen der Länge von der Größenordnung 2^N auftreten. (Die sehr langen Zyklen sind bei $g = 1$ wahrscheinlicher als bei größeren g .) Die Abb. 5.8 zeigt relevante numerische Simulationen, in denen sehr lange Zyklen für verschiedene Werte von N gefunden wurden. Die Existenz der sehr langen Zyklen ist wohl die wichtigste Erkenntnis aus der Betrachtung der einfacheren Komponenten der booleschen Netzwerke.

Das Auftreten von besonders langen Perioden in deterministischen diskreten dynamischen Systemen ist in der Mathematik seit geraumer Zeit bekannt und fand beispielsweise Anwendung in einer Klasse der Zufallszahlengeneratoren, in den so genannten *Additive Lagged Fibonacci Generators* [Mars 84]. In diesen Zufallsgeneratoren wird eine Folge von m -Bits - Zahlen x_k erzeugt gemäß

$$x_k = x_{k-p} + x_{k-p+q} \bmod m.$$

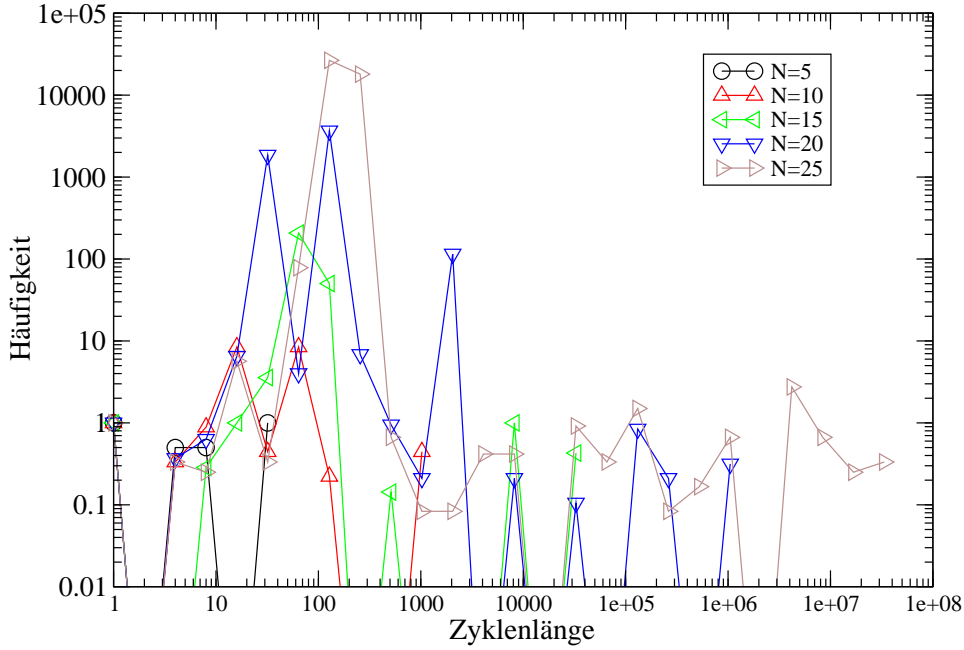


Abb. 5.8.: Anzahl der Zyklen mit Zykluslängen aus den Intervallen $[2^n, 2^{n+1})$ für die reversible Funktion $f_\Sigma = f_9$, gemittelt jeweils für ein gegebenes N über alle möglichen Werte von L .

Bei $m = 1$, $p = N$, $q = L$ ergibt diese Regel die Bitfolge der Knotenwerte von Σ im Fall der reversiblen Funktion $f_\Sigma = f_6$ in der Schleife mit einer Querverbindung.

Die mittlere Anzahl und Länge der Zyklen als Funktion von N ist in Abb. 5.9 dargestellt und wird im Folgenden erläutert. Die mittlere Zyklusanzahl

$$\bar{C}_N = \frac{1}{N-1} \sum_{L=1}^{N-1} C_{N,L} \quad (5.14)$$

wächst exponentiell mit N für die Komponentengrößen, die keine Primzahlen sind. Für die mittlere Zykluslänge \bar{P}_N gibt es mehrere Definitionsmöglichkeiten.

Erstens, als die Mittelung über alle Zyklen von allen Komponenten bzw. bei allen L ,

$$\bar{P}_N^{(1)} = \frac{\sum_L C_{N,L} \bar{P}_{N,L}}{\sum_L C_{N,L}}. \quad (5.15)$$

Mit dieser Definition gilt

$$\bar{P}_N^{(1)} \bar{C}_N = 2^N. \quad (5.16)$$

Diese Relation wird durch numerische Simulationsdaten im oberen Teil der Abb. 5.9 erfüllt. Die mittlere Zykluslänge hat lokale Maxima für die N , die Primzahlen sind. Gleichzeitig ist die mittlere Anzahl der Zyklen lokal am kleinsten.

5. Einfache Netzwerkkomponenten

Zweitens, als der Durchschnitt der mittleren Zykluslängen der Komponenten mit fixierten L ,

$$\bar{P}_N^{(2)} = \frac{1}{L} \bar{P}_{N,L}. \quad (5.17)$$

Diese Definition ist physikalisch sinnvoller, da jede Komponentenrealisierung gleich gewichtet werden sollte. Damit erhält man eine "glatte" exponentiell ansteigende Abhängigkeit der mittleren Zykluslänge von der Komponentengröße, siehe im unteren Teil der Abb. 5.9.

Eine dritte mögliche Definition der mittleren Zykluslänge, die jeden Anfangszustand der Komponenten gleich gewichtet, führt zu noch größeren Werten.

Ergänzende Ergebnisse

Im Anhang C wird ergänzend analytisch gezeigt, dass für alle Schleifen der Größe $N = 2^\nu N_0$ mit den ganzen Zahlen ν und N_0 die Ungleichung $C_N \geq 2^{AN}/N$ mit einer positiven Konstante A gilt, was auf eine exponentielle Abhängigkeit schließen lässt. Außerdem wird erklärt, warum in numerischen Simulationen jede Komponentenrealisierung durchschnittlich immer einen Fixpunkt besitzt und warum im Fall der kanalisierenden Funktionen mit Wahrscheinlichkeit $1/4$ ein Zyklus der Länge 2 auftritt.

5.2.3. Anwendbarkeit der Simulationen

In diesem Abschnitt wird argumentiert, dass Netzwerke mit bereits wenigen Verzweigungen (als komplexe Systeme) sich wegen der Größe des Zustandsraums 2^N nur eingeschränkt am Computer simulieren lassen.

Die Simulationen im Kapitel 5 verdeutlichen, warum es so schwierig ist, genaue Werte für die Anzahl und Länge der Attraktoren in numerischen Simulationen zum kritischen RBN - Modell zu bestimmen.

Bereits einfache Netzwerkkomponenten besitzen Zyklen, die ausgehend von einem im Grenzfall großer Netzwerke verschwindend kleinen Anteil der Anfangszustände des Netzwerks erreicht werden. Beispielsweise gibt es in der Komponente aus zwei Schleifen mit einer Zwischenverbindung für eine gerade erste Schleife (Knotenanzahl N_1), die sich in einem der zwei Fixpunkte befindet, $\sim 2^{N_2}/N_2$ Attraktoren. Sie werden nur ausgehend von einem Anteil $\sim 2^{-N_1}$ der Anfangszustände der Komponente erreicht.

Außerdem treten im Fall, dass N und L für die Schleife mit einer Querverbindung bzw. dass N_1 und N_2 für zwei verbundene Schleifen viele gemeinsame Teiler haben, besonders viele verschiedenartige Attraktoren auf. Durch eine Stichprobe (Sampling) von einem kleinen Teil der Anfangszustände der Komponenten ist es unmöglich all diese Zyklen zu treffen. Bei den Simulationen in diesem Kapitel 5 wurde deshalb immer die vollständige Suche im Zustandsraum durchgeführt.

Die vollständige Suche im Zustandsraum reicht auch nicht aus um zuverlässige Ergebnisse für das asymptotische Verhalten zu bekommen. Am Beispiel der Schlei-

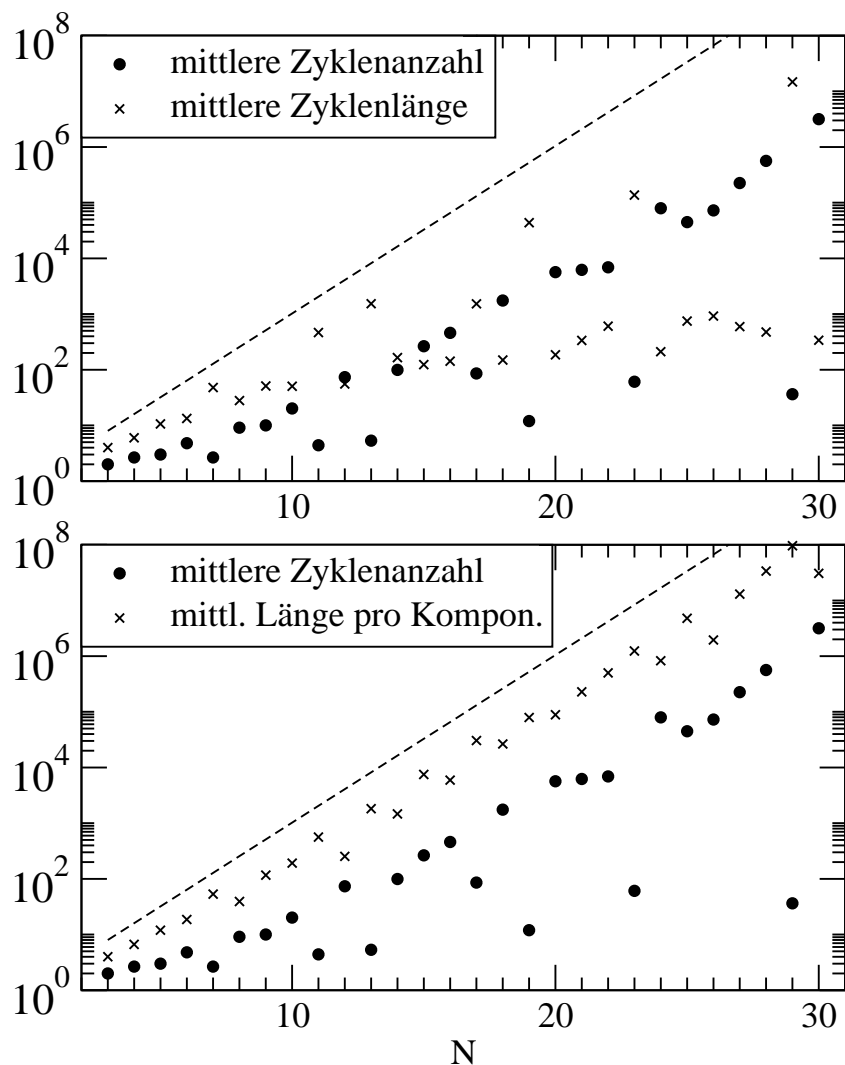


Abb. 5.9.: Mittlere Zyklusanzahl und Länge als Funktion von N . Oben: Die Zykluslängen werden nach (5.15) über alle Zyklen von allen Komponentenrealisierungen gemittelt. Unten: Es werden zuerst die mittleren Zykluslängen für jede Komponentenrealisierung separat bestimmt, dann wird nach (5.17) der Durchschnitt gebildet. Die gestrichelte Linie ist die Funktion 2^N .

5. Einfache Netzwerkkomponenten

fe mit einer Querverbindung und einer kanalisierenden Funktion des Knotens mit zwei Eingängen sieht man, dass die analytisch bekannte mittlere Anzahl der Zyklen bei den für die Simulation zugänglichen Komponentengrößen nicht erreicht werden kann. In diesem Beispiel skalieren die Beiträge der verschiedenen Knotenfunktionen und verschiedenen L unterschiedlich mit N , der asymptotisch größte Beitrag dominiert erst bei sehr großen N .

Eine weiterführende Diskussion der Einschränkungen beim Einsatz der numerischen Simulationen zur Untersuchung des RBN - Modells findet man im Abschn. 6.2, siehe auch die Einleitung zum Kapitel 6.

6. Kritische RBN mit $K = 2$

In diesem Kapitel werden dynamische Eigenschaften der RBN mit $K = 2$ untersucht. Nach Kapitel 4 und insbesondere Abschn. 4.4 haben diese Netzwerke, einschließlich der kanalisierenden Netzwerke und der zu den kritischen $K = 1$ Netzwerken analogen \mathcal{C}_1 Netzwerke, eine zueinander ähnliche Struktur, die durch wenige von der Wahl der Knotenfunktionen abhängige Parameter beschrieben werden kann. Die Dynamik der verschiedenen Netzwerke spiegelt diese Ähnlichkeit wider.

Die dynamischen Eigenschaften der Attraktoren numerisch zu simulieren erweist sich wegen großer statistischer Fluktuationen und der nach Abschn. 5.2.3 prinzipiellen Unzuverlässigkeit der Daten als schwierig. Eine Interpretation der Daten ist nur mit Hilfe der in der Literatur und in dieser Arbeit gewonnenen analytischen Ergebnisse möglich. Dieses Problem wird uns sowohl im Abschn. 6.1 beim Studium der gefrorenen Knoten als auch im Abschn. 6.2 bei der Betrachtung der mittleren Anzahl der Attraktoren und der Wahrscheinlichkeitsverteilung für die Länge der Attraktoren begegnen. Insbesondere stellt eine Nachprüfung der analytischen Ergebnisse in [Samu 03, Dros 05a] eine sehr komplizierte Aufgabe dar.

Im Abschn. 6.1 wird die in der Literatur wenig beachtete Frage untersucht, ob immer dieselben Knoten auf den Attraktoren gefroren sind. Es stellt sich heraus, dass die mittlere Anzahl der Knoten, die auf mindestens einem, aber nicht auf allen Attraktoren gefroren sind, genau so wie die mittlere Anzahl ($\sim N^{2/3}$) der nicht gefrorenen Knoten mit der Netzwerkgröße N skaliert. In \mathcal{C}_1 Netzwerken, die keine komplexen Komponenten haben, skaliert die mittlere Anzahl der “manchmal” gefrorenen Knoten dagegen nur wie die Anzahl der relevanten Knoten ($\sim \sqrt{N}$).

Ursprünglich wurden kritische RBN mit $K = 2$ als kritisch bezeichnet, weil man glaubte, dass die mittlere Anzahl der Attraktoren wie ein Potenzgesetz mit der Netzwerkgröße skaliert. Im Abschn. 6.2 wird auf eine anschauliche Weise gezeigt, dass die mittlere Anzahl und Länge der Attraktoren schneller als jedes Potenzgesetz ansteigt. Damit zeigen nicht die Attraktoren, sondern die nicht gefrorenen und die relevanten Knoten das Skalenverhalten.

In der Literatur wurde häufig behauptet [Kauf 93], dass die Attraktoren in den kanalisierenden \mathcal{C}_2 Netzwerken kürzer sind als in den übrigen kritischen Netzwerken. Im Abschn. 6.2.2 wird anschaulich gezeigt, dass in \mathcal{C}_2 Netzwerken die Attraktoren größenordnungsmäßig genau so lang wie die längsten Attraktoren in den allgemeinen RBN mit $K = 2$ sein können. Sie treten nur seltener als in den allgemeinen RBN auf. Im biologischen Kontext sind die kanalisierenden Netzwerke besonders interessant, siehe Abschn. 2.3.2.

6.1. Nicht auf allen Attraktoren gefrorene Knoten

Für große kanalisierende \mathcal{C}_2 Netzwerke mit N Knoten wurde im Abschn. 4.2.2 gezeigt, dass der gefrorene Kern alle außer die größenordnungsmäßig $N^{2/3}$ auf mindestens einem Attraktor nicht gefrorenen Knoten enthält. Das stimmt mit dem analytischen Ergebnis aus [Samu 03, Dros 05a] überein, dass die mittlere Anzahl der nicht gefrorenen Knoten mit $N^{2/3}$ skaliert. Allerdings sagt dieses Ergebnis nichts darüber aus, ob die gefrorenen Knoten auf allen Attraktoren gefroren sind und damit eine strukturelle Klasse von Knoten definieren oder ob sie “dynamisch” auf manchen Attraktoren einfrieren und auf anderen nicht¹. Diese Frage kann erst mit dem Strukturwissen aus dem Abschn. 4.4 beantwortet werden. Zunächst sollen numerische Simulationen zeigen, welche Knoten gefroren sind. Danach werden die Ergebnisse erklärt.

In Abb. 6.1 ist die mittlere Anzahl der gefrorenen Knoten $N_f^{(a)}$ pro Attraktor für die kanalisierenden Netzwerke und für sRBN dargestellt. Genauer ist dort die Differenz $(N - N_f^{(a)})$ als Funktion von N aufgetragen um das Verhalten im thermodynamischen Grenzfall $N \rightarrow \infty$ studieren zu können. Die anderen zwei Kurven zeigen die Differenz $(N - N_f^{(n)})$, wobei $N_f^{(n)}$ die mittlere Anzahl der auf allen Attraktoren gefrorenen Knoten ist. Es können nur die Attraktoren, die in der jeweiligen Simulation gefunden wurden, berücksichtigt werden, weitere technische Details zu den Simulationen können der Bildunterschrift entnommen werden.

Die Form der Kurven zur mittleren Anzahl der nicht gefrorenen Knoten pro Attraktor $(N - N_f^{(a)})$ ist fast identisch für sRBN und für \mathcal{C}_2 Netzwerke, in Übereinstimmung mit den Aussagen des Abschn. 4.4 über das relevante Teilnetzwerk in verschiedenen Klassen von RBN mit $K = 2$ bzw. in Übereinstimmung mit den Ergebnissen in [Samu 03, Dros 05a]. Das erwartete asymptotische Potenzgesetz $N^{2/3}$ wurde in den Simulationen wegen der sehr langsamen asymptotischen Konvergenz jedoch nicht erreicht. (Die Anzahl der relevanten Knoten war in den Simulationen nur von der Größenordnung 10, die Anzahl der nicht gefrorenen Knoten von der Größenordnung 100 für die größten betrachteten Netzwerke.) Das wird für die Diskussion weiter unten nicht ausschlaggebend sein.

Ich gehe davon aus, dass sich $(N - N_f^{(a)})$ in den beiden Klassen von Netzwerken, \mathcal{C}_2 und sRBN, ähnlich verhält. Für die Kurven zu $(N - N_f^{(n)})$ findet man in den Simulationen, dass sie sich jeweils nur durch einen konstanten Faktor von den Kurven zu $(N - N_f^{(a)})$ für beide Netzwerktypen unterscheiden. Somit implizieren die Simulationen, dass die kanalisierenden Netzwerke einen gefrorenen Kern (beschrieben durch $N_f^{(n)}$) und größenordnungsmäßig $N^{2/3}$ nicht gefrorene Knoten haben. Zusätzlich gibt es größenordnungsmäßig $N^{2/3}$ Knoten, die nur auf einem Teil der Attraktoren gefroren sind. Der angesprochene Faktor ist für die kanalisierenden Netzwerke größer als für sRBN.

¹Ohne die Ergebnisse im Abschn. 4.2.2 könnte man aufgrund der Abwesenheit der gefrorenen Funktionen in den kanalisierenden Netzwerken vermuten, dass \mathcal{C}_2 Netzwerke die nur auf manchen Attraktoren gefrorene Knoten beinhalten und keinen gefrorenen Kern haben.

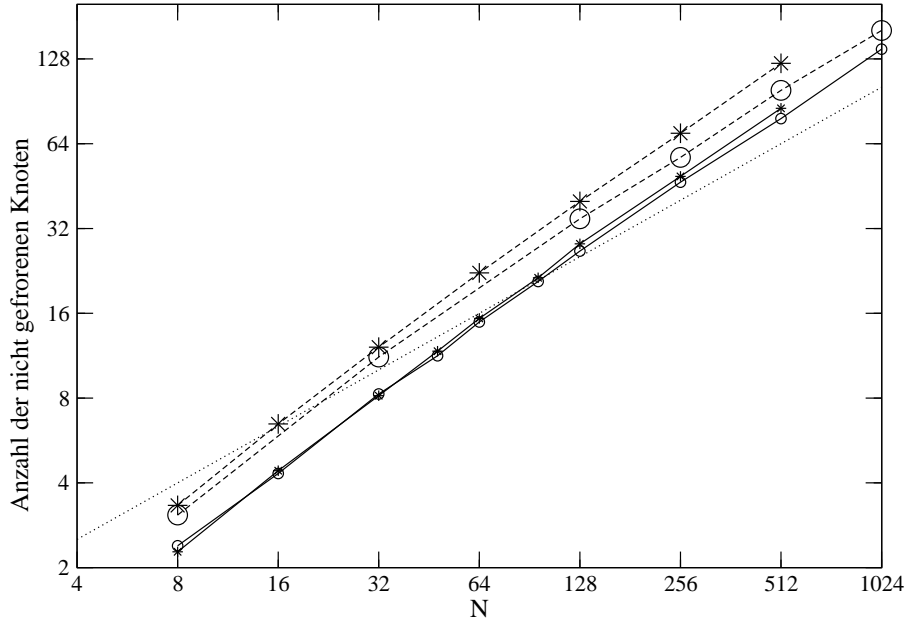


Abb. 6.1.: Mittlere Anzahl der nicht gefrorenen Knoten für kanalisierende \mathcal{C}_2 Netzwerke (*) und (stellvertretend für sonstige RBN mit $K = 2$) sRBN (o). Die durchgezogenen Linien verbinden die Datenpunkte für $N - N_f^{(a)}$ (die mittlere Anzahl der nicht gefrorenen Knoten pro Attraktor), die gestrichelten Linien verbinden die Datenpunkte für $N - N_f^{(n)}$ (die mittlere Anzahl der Knoten, die mindestens auf einem gefundenen Attraktor nicht gefroren sind). Die punktierte Gerade steht für das Potenzgesetz $N^{2/3}$. Die verschiedenen Attraktoren werden nur einmal ohne Berücksichtigung der Größe des Einzugsbereichs gezählt. Die Daten wurden für 1000 Anfangszustände pro Netzwerk ausgewertet und über 2000 Netzwerke gemittelt.

Warum gibt es einen konstanten Faktor, der die Größen $(N - N_f^{(a)})$ und $(N - N_f^{(n)})$ verbindet? Der Unterschied zwischen den Kurven $(N - N_f^{(a)})$ und $(N - N_f^{(n)})$ wird durch die Knoten verursacht, die auf mindestens einem, aber nicht auf allen Attraktoren gefroren sind. Um den Faktor zu erklären muss man daher zeigen, dass ein Anteil ungleich Null der nicht (auf allen Attraktoren) gefrorenen Knoten auf einem Anteil ungleich Null aller Attraktoren sich auf einem Attraktor der Länge 1, also in einem Fixpunkt, befindet.

Die Knoten, die nur auf manchen Attraktoren gefroren sind, gehören nicht zum gefrorenen Kern. Sie sind nach Abschn. 4.4 entweder relevant oder sie sitzen auf nicht gefrorenen in den relevanten Komponenten wurzelnden Bäumen. Wir betrachten zuerst die häufigsten und einfachsten relevanten Komponenten, die einfachen Schleifen, und ihre nicht gefrorenen Bäume. Die mittlere Anzahl der relevanten Schleifen der Länge l ist $1/l$ für alle l bis zur Abschneidelänge $l_c \sim N^{1/3}$, die mittlere Größe der Bäume ist $N^{1/3}$. Die größte nicht gefrorene Komponente, deren

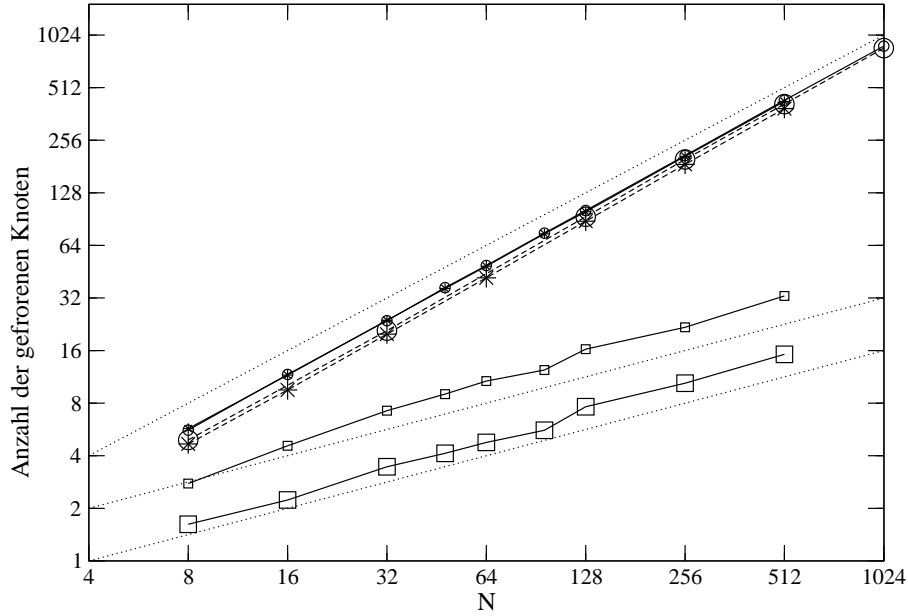


Abb. 6.2.: Mittlere Anzahl der gefrorenen Knoten für \mathcal{C}_1 Netzwerke. Große \square stehen für die Werte von $N_f^{(n)}$, kleine \square stehen für $N_f^{(a)}$. Zum Vergleich sind die entsprechenden Kurven für sRBN und \mathcal{C}_2 Netzwerke eingezeichnet, die Daten stammen aus der Abb. 6.1. Die punktierten Geraden stehen für die Potenzgesetze N , \sqrt{N} und $\sqrt{N}/2$.

relevante Knoten auf einer einfachen Schleife sitzen, hat unter Berücksichtigung der Bäume also von der Größenordnung $N^{2/3}$ Knoten, einen Anteil ungleich Null aller nicht gefrorenen Knoten. Wenn die einfache Schleife in einem Fixpunkt ist, dann friert die gesamte Komponente ein. Da diese Komponente im Limes $N \rightarrow \infty$ nach Abschn. 4.5.1 mit einer endlichen Wahrscheinlichkeit vorkommt, reicht es zur Erklärung des obigen Faktors zu zeigen, dass die relevante Schleife einen Fixpunkt mit einer endlichen Wahrscheinlichkeit erreicht.

Eine gerade relevante Schleife der Länge l hat zwei Fixpunkte (siehe Abschn. 5.1), eine ungerade hat keinen. Jeder der zwei Fälle kommt mit Wahrscheinlichkeit $1/2$ vor. Die Anzahl der Attraktoren wächst jedoch exponentiell mit l ($\sim 2^l/l$). Aus diesem Grund hat, für große N , nur ein verschwindender Anteil aller Attraktoren der Komponente der Größe l_c die Länge 1.

Als Nächstes betrachten wir die wenigen komplexen relevanten Komponenten mit Knoten mit zwei relevanten Eingängen (vgl. wieder mit dem Abschn. 4.4). Im Unterschied zu den einfachen Schleifen, wo jeder Anfangszustand Teil eines Zyklus im Zustandsraum ist, können komplexe relevante Komponenten attraktive Fixpunkte mit einem großen Einzugsbereich besitzen, der einen nicht verschwindenden Anteil kleiner Eins des Zustandsraumes der Komponente ausmacht. Ein Beispiel ist die im Abschn. 5.2.2 untersuchte Schleife mit einer Querverbindung und der Funktion f_{14} an der Verzweigung. Da diese Komponente mit einer endlichen Wahrscheinlichkeit

im Netzwerkensemble vorkommt, bekommt der Faktor zwischen den Kurven für $(N - N_f^{(n)})$ und für $(N - N_f^{(a)})$ in Abb. 6.1 eine Erklärung. Zusätzlich wird klar, warum dieser Faktor für \mathcal{C}_2 Netzwerke größer als für sRBN ist. Die Wahrscheinlichkeit, dass der Verzweigungsknoten im obigen Beispiel die spezielle Funktion f_{14} hat ist größer für die kanalisierenden Netzwerke.

Der Vollständigkeit halber untersuche ich noch die manchmal gefrorenen Knoten in \mathcal{C}_1 Netzwerken. Die Simulationsergebnisse für $N_f^{(a)}$ und $N_f^{(n)}$ in diesen Netzwerken sind in Abb. 6.2 dargestellt. Beide Kurven steigen wie \sqrt{N} mit der Netzwerkgröße an. Die Größe des gefrorenen Kerns $N_f^{(n)}$ beträgt asymptotisch annähernd $\sqrt{N}/2$, weil nur gerade relevante Schleifen der Länge 1 auf allen Attraktoren gefroren sind. Dabei beträgt die mittlere Größe der Bäume nach Abschn. 4.4 \sqrt{N} , $1/2$ ist die Wahrscheinlichkeit, dass eine Schleife der Länge 1 gerade ist.

Als Nächstes wird $N_f^{(a)}$ bestimmt. Nur gerade Schleifen der Länge l haben zwei Fixpunkte, diese Fixpunkte werden für $2^l - 2$ Anfangszustände nicht erreicht. Der Beitrag von den einfachen Schleifen der Länge l zur mittleren Anzahl der gefrorenen Knoten pro Attraktor ist $1/l \times 1/2 \times 2l/2^l \times \sqrt{N}$, wobei $1/2$ die Wahrscheinlichkeit für eine gerade Schleife und \sqrt{N} die mittlere Baumgröße ist. Summation über l ergibt $N_f^{(a)} \simeq \sqrt{N}$. Folglich beträgt das Verhältnis $(N - N_f^{(n)})/(N - N_f^{(a)})$ asymptotisch 1 im Unterschied zu \mathcal{C}_2 Netzwerken. Die Begründung liegt darin, dass \mathcal{C}_1 Netzwerke einfache relevante Schleifen und keine komplexen relevanten Komponenten haben.

6.2. Anzahl und Länge der Attraktoren

Die relevanten Knoten bestimmen nach Definition die Dynamik im RBN - Modell, insbesondere die mittlere Anzahl und Länge der Attraktoren. Wir betrachten kritische Netzwerke mit N Knoten und $K = 2$ Eingängen pro Knoten im thermodynamischen Limes $N \rightarrow \infty$ bzw. für genügend große N . Die Anzahl der relevanten Knoten in diesen Netzwerken N_{rel} beträgt nach Abschn. 4.4 größenordnungsmäßig $N^{1/3}$. Die relevanten Knoten sind in unabhängige einfache Schleifen und komplexe relevante Komponenten organisiert.

- Es gibt nach Abschn. 4.5.1 meistens nicht mehr als eine komplexe relevante Komponente pro Netzwerk mit größenordnungsmäßig N_{rel} Knoten. Sie enthält dann alle relevanten Knoten mit zwei relevanten Eingängen; ihre Anzahl ist konstant unabhängig von der Netzwerkgröße.
- Alle anderen von der Größenordnung $\ln(N_{rel})$ Komponenten sind einfache Schleifen aus relevanten Knoten mit effektiv einem Eingang. Für die Abschneidelänge dieser Schleifen gilt $l_c \sim N_{rel}$. Einfache Schleifen mit Längen bis zur Abschneidelänge sind analog zu denen des relevanten Teilnetzwerks der kritischen $K = 1$ Netzwerke. (In den kritischen $K = 1$ Netzwerken mit n Knoten gibt es größenordnungsmäßig \sqrt{n} relevante Knoten, die Abschneidelänge skaliert wieder mit der Anzahl der relevanten Knoten.)

Wenn wir nur die einfachen relevanten Schleifen im Netzwerkensemble betrachten, dann ist die entsprechende mittlere Anzahl und mittlere Länge der Attraktoren eine untere Schranke für diese Größen im ursprünglichen Netzwerkensemble. Die einfachen Schleifen entsprechen aber dem kritischen $K = 1$ Netzwerkensemble, wo die mittlere Anzahl und Länge der Attraktoren bekanntlich [Dros 05b] schneller als jede Potenz mit der Netzwerkgröße ansteigt.

Folglich steigt die mittlere Anzahl und Länge der Attraktoren in kritischen $K = 2$ Netzwerken schneller als jede Potenz mit der Netzwerkgröße an. Zur selben Schlussfolgerung kommt man durch Betrachtung der komplexen relevanten Komponenten, da jede von ihnen nach Abschn. 4.5.1 und 3.2.1 mit einer Wahrscheinlichkeit ungleich Null im Netzwerkensemble vorkommt und da es nach Abschn. 5.2.2 zumindest eine komplexe relevante Komponente gibt, die für sich genommen eine exponentiell mit der Komponentengröße wachsende mittlere Anzahl und Länge der Attraktoren besitzt. Für die mittlere Anzahl der Attraktoren ergibt sich der historisch erste analytische Beweis (vgl. Abschn. 2.4) aus der Analyse der Anzahl der Attraktoren einer vorgegebenen Länge L , siehe [Samu 03].

Das obige Resultat für die mittlere Anzahl und Länge der Attraktoren wurde erst vor wenigen Jahren bewiesen, weil die Daten bei früheren numerischen Simulationen (vgl. Abschn. 2.4) fehlinterpretiert wurden. Entweder wurden nur kleinere Netzwerke betrachtet, denen das asymptotische Verhalten nicht anzusehen war, oder die prinzipiellen Einschränkungen der numerischen Simulationen aufgrund des Undersampling² wurden unterschätzt. Allgemein stoßen numerische Simulationen aus folgenden Gründen an ihre Grenzen, siehe auch Abschn. 5.2.3.

- Das asymptotische Verhalten stellt sich erst bei einer genügend großen Anzahl der relevanten Knoten ein. Da diese Anzahl mit der Netzwerkgröße sehr langsam ansteigt (wie $N^{1/3}$), sind die in Simulationen zugänglichen Netzwerkgrößen derzeit meist zu klein um das asymptotische Verhalten zu sehen.
- Undersampling äußert sich darin, dass sehr kurze und sehr lange Attraktoren übersehen werden. Die ersteren machen einen exponentiell kleinen Anteil des durch die Netzwerkgröße N charakterisierten Zustandsraums aus, die letzteren erscheinen exponentiell selten in einer Netzwerkrealisierung.

Ich greife die Herausforderung der numerischen Simulation großer Netzwerke wieder auf und untersuche, ausgestattet mit dem Wissen über die Struktur der Netzwerke aus dem Kapitel 4 und mit den analytischen Ergebnissen aus [Samu 03, Dros 05a], wo die Grenzen der Anwendbarkeit liegen und inwiefern Interpretationen zulässig sind. Dabei ist die beispielsweise im Abschn. 4.5.1 verwendete effiziente Methode zur Netzwerksimulation mit Hilfe der stochastischen Prozesse aus dem Ka-

²Der Zustandsraum wächst exponentiell und der Netzwerkkonfigurationsraum wächst überexponentiell mit der Netzwerkgröße. Numerisch kann nur ein zunehmend kleiner Anteil aller Anfangsbedingungen und aller Netzwerkkonfigurationen untersucht werden, was zu Fehlern (Undersampling) führt.

pitel 4 wenig nützlich, da die meiste Computerzeit bei der Suche der Attraktoren verloren geht.

Es werden im Folgenden statistische Eigenschaften der Attraktoren

- der kritischen $K = 1$ Netzwerke (\mathcal{C}_1 Netzwerke),
- der kanalisierenden $K = 2$ Netzwerke (\mathcal{C}_2 Netzwerke) und
- der sRBN als Repräsentanten der Klasse der allgemeinen RBN mit $K = 2$

in numerischen Simulationen für große Netzwerkgrößen untersucht und interpretiert. Im Abschn. 6.2.1 wird ein Versuch unternommen, analytisch bekannte Ergebnisse für die mittlere Anzahl der Attraktoren einer vorgegebenen Länge numerisch zu bestätigen. Im Abschn. 6.2.2 wende ich mich der Wahrscheinlichkeitsverteilung der Attraktorlängen und der mittleren Attraktorlänge zu. Im Abschn. 6.2.3 wird die mittlere Anzahl der Attraktoren diskutiert.

6.2.1. Mittlere Anzahl der Attraktoren der Länge L

Für sRBN wurden dynamische Attraktoren direkt in der bemerkenswerten analytischen Arbeit [Samu 03] untersucht. In [Dros 05a] wurden die Ergebnisse für sRBN auf allgemeine kritische RBN mit $K = 1$ und $K = 2$ ausgeweitet und es wurden Brücken zu dieser Arbeit (insbesondere zum Abschn. 4.4) geschlagen. Die mittlere Anzahl $C_L(N_{rel})$ der Attraktoren der Länge L skaliert demnach wie eine Potenz der Anzahl der relevanten Knoten,

$$C_L(N_{rel}) \sim N_{rel}^{G_L - 1}, \quad (6.1)$$

mit einem von L abhängenden Vorfaktor. G_L hängt mit der Anzahl der Zyklen einer einfachen Schleife der Länge L zusammen:

$$G_L \sim 2^L / L, \quad (6.2)$$

die Details findet man am Ende des zweiten Abschnitts in [Dros 05a].

In Abb. 6.3 sind die Simulationsergebnisse für $C_7(N_{rel})$ und $C_8(N_{rel})$ jeweils für \mathcal{C}_1 und \mathcal{C}_2 Netzwerke dargestellt. Ohne zusätzliche Informationen wären die Daten nutzlos. Bei vorsichtiger Interpretation widersprechen sie aber, wie ich gleich zeige, nicht der theoretischen Skalierung (6.1).

Die betrachteten 1000 Anfangszustände pro Netzwerkrealisierung entsprechen dem Zustandsraum eines Netzwerks mit nicht mehr als etwa 10 relevanten Knoten. Das erklärt, warum für $N_{rel} \gtrsim 10$ die Anzahl der gefundenen Attraktoren entgegen der analytischen Formel mit N_{rel} abnimmt. Andererseits gilt die Formel (6.1) nur für $N_{rel} \gg L$, sodass es derzeit unmöglich erscheint, das Skalenverhalten (6.1) in Computersimulationen zu sehen. Trotzdem enthalten die Daten in Abb. 6.3 nützliche Informationen. Der auffallende qualitative Unterschied der Kurven für gerade und ungerade L (auch sichtbar für andere L , nicht dargestellt) entspricht

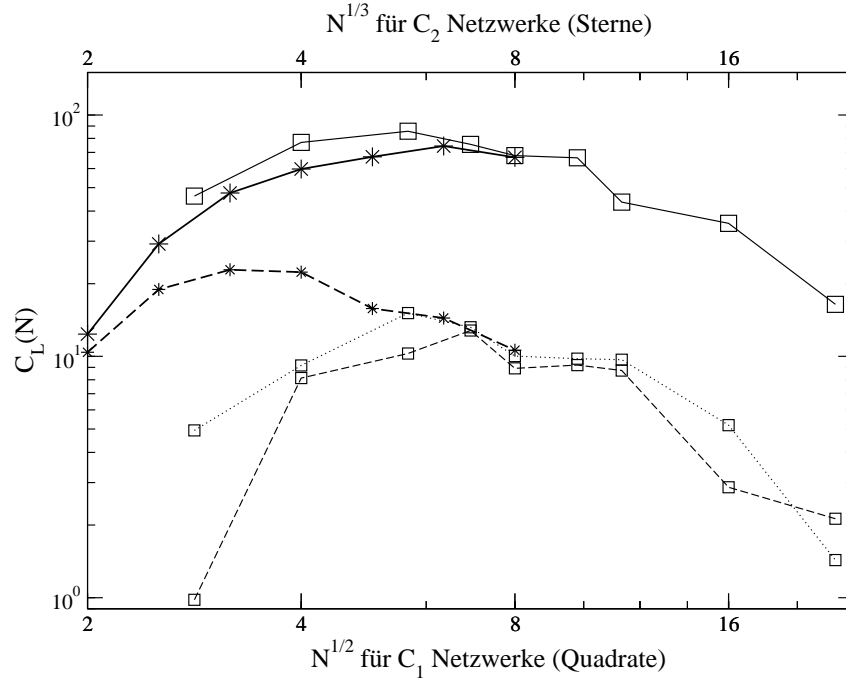


Abb. 6.3.: Die absolute Anzahl der verschiedenen gefundenen Attraktoren der Längen $L = 7$ (gestrichelte und punktierte Linien) und $L = 8$ (durchgezogene Linien) für kanalisierende \mathcal{C}_2 (*) und \mathcal{C}_1 (\square) Netzwerke als Funktion der Netzwerkgröße. Die Abszissen der Kurven für die beiden Klassen von Netzwerken geben jeweils größenordnungsmäßig die Anzahl der relevanten Knoten an. Die Datenpunkte wurden über 2000 Netzwerkrealisierungen gemittelt, für jede von ihnen wurden 1000 Anfangszustände betrachtet. Zum Zweck des Vergleichs entspricht die punktierte Kurve der Mittelung über 1000 und nicht 2000 Netzwerke.

dem Ergebnis aus [Dros 05a], dass G_L kleiner für ungerade L als für die an sie angrenzenden geraden L ist. Außerdem ist der Verlauf der Kurven für \mathcal{C}_1 und \mathcal{C}_2 Netzwerke ähnlich, in Übereinstimmung mit der Erkenntnis aus dem Abschn. 4.4, dass die relevanten Teilnetzwerke in den beiden Klassen von Netzwerken gleiche topologische Eigenschaften für die meisten relevanten Komponenten haben.

6.2.2. Verteilung der Attraktorlängen

Wir fangen an mit der Auswertung der in Abb. 6.4 dargestellten Wahrscheinlichkeitsverteilung der Attraktorlängen. Dabei sind die bereits bei der Diskussion der Simulationsergebnisse in Abb. 6.3 verwendeten Formeln (6.1) und (6.1) nützlich. Asymptotisch für $N \rightarrow \infty$ erwartet man mit (6.1), dass die Wahrscheinlichkeiten in Abb. 6.4 für steigende L zunehmen (für die im Vergleich zur Anzahl der relevanten Knoten nicht zu großen Attraktorlängen L). Das ist in den Simulationen nicht der Fall, was auf einen Effekt endlicher Systemgröße (“finite size effect”) hindeutet.

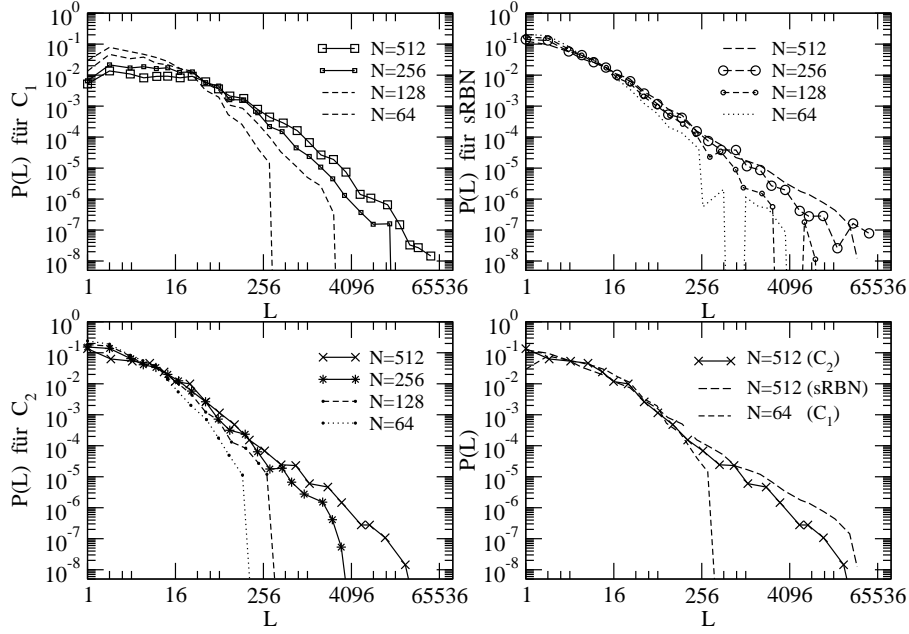


Abb. 6.4.: Wahrscheinlichkeitsverteilung für die Attraktorlängen, aufgetragen für verschiedene konstante Netzwerkgrößen N . Die vier Graphen stellen die Ergebnisse für C_1 Netzwerke, sRBN (stellvertretend für kritische RBN mit $K = 2$), kanalisierende C_2 Netzwerke und, rechts unten, ein Vergleich der drei Klassen von Netzwerken dar. Bei diesem Vergleich haben C_1 Netzwerke mit 64 Knoten im Mittel die gleiche Anzahl der relevanten Knoten wie C_2 Netzwerke mit 512 Knoten. Für 2000 Netzwerkrealisierungen wurden für 1000 Anfangszustände pro Netzwerkrealisierung die verschiedenen Attraktoren bestimmt. Die Attraktorlängen wurden mit der Größe ihres Einzugsbereichs gewichtet und in Intervallen auf einer logarithmischen Skala zusammengefasst. Aufgetragen ist jeweils die so ermittelte relative Häufigkeit der Attraktorlängen in Intervallen. Die Daten, die hier und in Abb. 6.3 ausgewertet wurden, stammen aus denselben Simulationen.

In der rechten unteren Ecke der Abb. 6.4 stimmen die Kurven für sRBN und C_2 Netzwerke bei $N = 512$ gut mit der Kurve für C_1 Netzwerke bei $N = 64$ überein, insbesondere bei kleineren L . Das ist nach Abschn. 4.4 nicht verwunderlich, da in diesem Beispiel alle drei Klassen von Netzwerken ungefähr die gleiche Anzahl der relevanten Knoten haben. Die Anzahl der relevanten Knoten N_{rel} skaliert nämlich mit $N^{1/3}$ für sRBN und C_2 und mit \sqrt{N} für C_1 Netzwerke, es gilt $512^{1/3} = 64^{1/2}$; die größenordnungsmäßig kleineren als $\sim N_{rel}$ relevanten Komponenten sind in jedem Fall einfache Schleifen; die mittlere Anzahl der einfachen Schleifen der Länge l ist $1/l$. Die Abweichung der Kurve für C_1 Netzwerke bei größeren L kann durch die Anwesenheit der komplexen relevanten Komponenten in sRBN und C_2 Netzwerken erklärt werden. In C_1 Netzwerken kommen ausschließlich einfache relevante Schleifen vor, ihre Zykluslängen wachsen linear mit der Größe der Schleifen. Dagegen können komplexe relevante Komponenten viel längere Attraktoren haben.

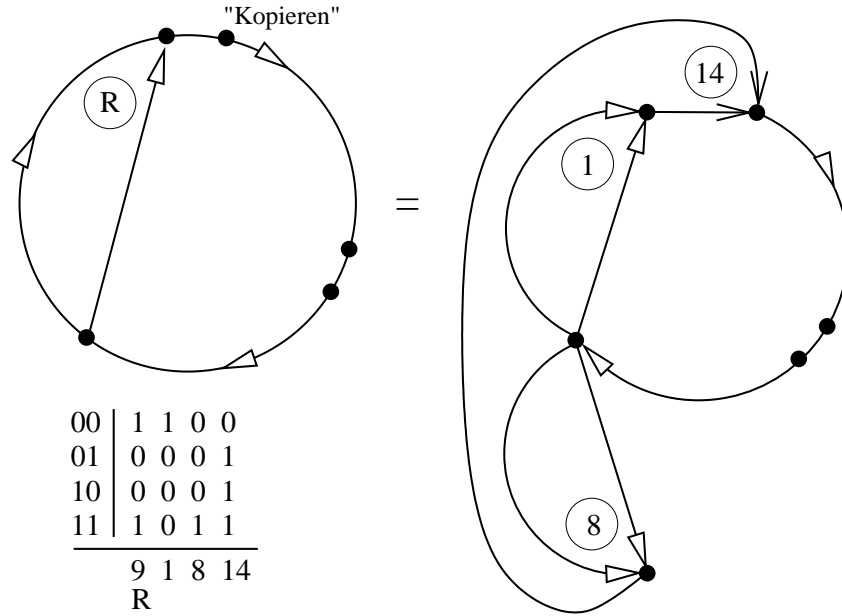


Abb. 6.5.: Eine relevante Komponente mit einer reversiblen Funktion und eine relevante Komponente aus einem \mathcal{C}_2 Netzwerk, deren Attraktoren paarweise aufeinander abgebildet werden können. Die nicht ausgefüllten dreieckigen Pfeile repräsentieren Ketten von Knoten, die für die beiden oberen Schleifen mit Querverbindung paarweise aufeinander abzubilden sind. Die übrigen Pfeile sind einzelne Verbindungen der Knoten. Ein Paar Knoten sind exemplarisch durch ausgefüllte Kreise dargestellt. Die untersten zwei Knotenkette in der rechten Komponente, die runde und die gerade, sind von ihrem Ursprung aus gesehen identisch mit den aus diesem Ursprung ausgehenden anderen zwei Knotenkette, der runden und der geraden. Die Wahrheitstabelle erläutert die Wahl der Funktionen, der reversiblen Funktion (9) und der drei kanalisierenden Funktionen (1, 8 und 14), an den Verzweigungen. Am Knoten mit der Funktion 14 kommt der boolesche Eingangszustand 11 niemals vor.

Schließlich verursachen die in \mathcal{C}_2 Netzwerken abwesenden reversiblen Funktionen den kleinen Unterschied zwischen den Kurven für sRBN und \mathcal{C}_2 Netzwerke bei größeren L . Aus Kapitel 5 wissen wir nämlich, dass komplexe relevante Komponenten mit Knoten mit reversiblen Funktionen und mit zwei relevanten Eingängen sehr große Attraktorlängen von der Größenordnung des Zustandsraums der Komponenten haben können.

Bemerkenswert, sehr lange Attraktoren wie in sRBN treten auch in \mathcal{C}_2 Netzwerken auf. Die \mathcal{C}_2 Netzwerke bekamen ursprünglich viel Aufmerksamkeit wegen ihrer vergleichsweise kurzen Attraktoren [Kauf 93]. Das Auftreten von sehr langen Attraktoren in kanalisierenden \mathcal{C}_2 Netzwerken erklären wir anhand eines Beispiels.

Wir betrachten die in sRBN vorkommende relevante Komponente in Form einer Schleife mit einer Querverbindung, die einen Knoten mit einer bestimmten reversiblen Funktion ihrer zwei relevanten Eingänge enthält. Nach Abschn. 5.2.2 können

die Attraktoren dieser Komponente einen endlichen Anteil des Zustandsraums sogar für sehr große Komponenten abdecken. Wir bauen als Nächstes eine kanalisierende relevante Komponente mit drei kanalisierenden \mathcal{C}_2 - Knotenfunktionen zusammen, die genau dieselben Attraktorzustände wie die Schleife mit Querverbindung hat. Die Abbildung der kanalisierenden Komponente auf die Schleife mit Querverbindung ist in Abb. 6.5 dargestellt. Die beiden Komponenten haben dieselbe Anzahl und dieselben Längen der Attraktoren. Allerdings kommt die durch die drei kanalisierenden Funktionen nachgebaute reversible Funktion nicht so häufig in \mathcal{C}_2 Netzwerken vor wie die reversiblen Funktionen in sRBN. Die sehr langen Attraktoren kommen in den kanalisierenden \mathcal{C}_2 Netzwerken daher vergleichsweise selten vor.

Wir kehren zurück zur Diskussion der Abb. 6.4. Um die ziemlich großen statistischen Fluktuationen zu glätten wurden die Attraktorlängen in genügend großen logarithmischen Intervallen zusammengefasst. Sonst würde man Schwankungen der Anzahl der Attraktoren mit benachbarten Längen zwischen geraden und ungeraden Längen sehen. Man erhält mehr Attraktoren für gerade Attraktorlängen als für benachbarte ungerade. Die kleineren relevanten Komponenten in RBN mit $K = 2$ sind einfache Schleifen. In ihnen kommen gerade Attraktorlängen häufiger vor, da eine gerade (ungerade) Schleife mit N_l Knoten Attraktoren der Länge N_l ($2N_l$) hat, sodass die geraden Attraktorlängen $2N_l$ sowohl in den Schleifen mit $2N_l$ als auch in den Schleifen mit N_l Knoten auftreten.

Die längsten gefundenen Attraktoren treten mit einer Wahrscheinlichkeit auf, die mit der Attraktorlänge exponentiell abnimmt. Das passiert ab einer gewissen Abschneidelänge, die sich phänomenologisch zu $L \sim AN^2$, mit $A = 0.1$ für sRBN und 0.01 für \mathcal{C}_2 Netzwerke, ergibt. Die in meinen Simulationen gefundenen Netzwerkrealisierungen mit untypisch langen Attraktoren können übrigens aufgrund von Undersampling zu verfälschten Ergebnissen für Ensemblemittelwerte wie die mittlere Attraktorlänge führen, vgl. Abb. 6.6.

Abschließend werden in Abb. 6.6 die Simulationsergebnisse für die mittlere Attraktorlänge vorgestellt. Die wichtigste Beobachtung dabei ist, dass die mittlere Attraktorlänge für die betrachteten Klassen von Netzwerken schneller als jedes Potenzgesetz mit der Netzwerkgröße N ansteigt. Insbesondere sind die Attraktoren in großen kanalisierenden Netzwerken nicht wesentlich kürzer als in den anderen Netzwerken. Nur bei kleinen N zeigen die Kurven und insbesondere diejenige für kanalisierende \mathcal{C}_2 Netzwerke ein \sqrt{N} Verhalten, das über viele Jahre hinweg angenommen wurde [Kauf 69]. Alle Kurven haben eine ähnliche Form, weil die entsprechenden relevanten Teilnetzwerke ähnlich sind. Die mittlere Attraktorlänge ist für große sRBN - Netzwerke, deren komplexe relevante Komponenten reversible Knotenfunktionen haben können, größer als für die anderen Klassen von Netzwerken.

6.2.3. Mittlere Anzahl der Attraktoren

Die numerischen Ergebnisse für die mittlere Anzahl der Attraktoren, die ausgehend von 1000 Anfangszuständen pro Netzwerkrealisierung gefunden wurden, sind in Abb. 6.7 dargestellt. Die Abszisse gibt für jede der Kurven größenordnungsma-

6. Kritische RBN mit $K = 2$

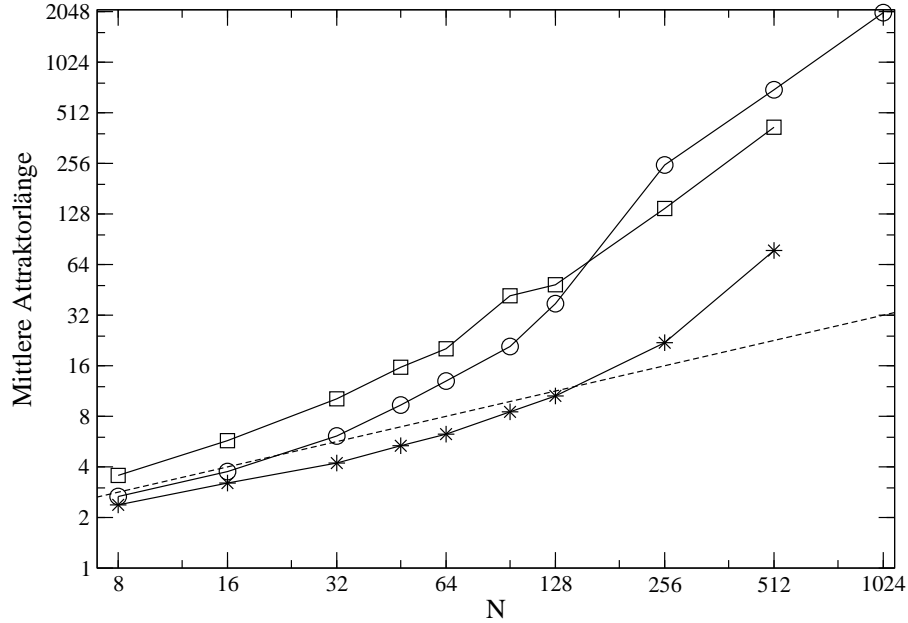


Abb. 6.6.: Mittlere Attraktorlänge für kanalisierende \mathcal{C}_2 Netzwerke (*), sRBN (o) und \mathcal{C}_1 Netzwerke (□) als Funktion der Netzwerkgröße N . Die punktierte Linie stellt das Potenzgesetz \sqrt{N} dar. Die einzelnen Attraktorlängen wurden jeweils mit der Größe des Einzugsbereichs gewichtet, es wurde über jeweils 1000 Anfangszustände für 2000 Netzwerkrealisierungen gemittelt.

big die entsprechende Anzahl der relevanten Knoten an. Bei den benutzten $\sim 2^{10}$ zufälligen Anfangszuständen nehmen die Knoten eines Netzwerks mit $N_{rel} \simeq 10$ relevanten Knoten einen großen Anteil der möglichen Werte an, sodass die mittlere Anzahl der gefundenen Attraktoren eine gute Abschätzung des genauen Ensemble-mittelwerts ist. Der kleine Unterschied zwischen den gestrichelten und den durchgezogenen Kurven in Abb. 6.7 zeigt, dass genügend viele Netzwerkrealisierungen für die Mittelung benutzt wurden. Für Netzwerke mit $N_{rel} \gg 10$ ist es unwahrscheinlich einen und denselben Attraktor mit den zufällig gewählten 1000 Anfangszuständen zweimal zu treffen. Aus diesem Grund erreicht die mittlere Anzahl der Attraktoren trivialerweise den Wert 1000 und enthält keine Informationen mehr über die Dynamik im System.

Die Ergebnisse in Abb. 6.7 wurden mit den numerischen Ergebnissen für den mittleren Medianwert der Anzahl der Attraktoren aus [Soco 03] verglichen. Die Medianwerte sind in den betrachteten Netzwerken kleiner als die mittlere Anzahl der Attraktoren und die entsprechende Kurve liegt unter den Kurven in Abb. 6.7 (nicht dargestellt). In [Dros 05b] wurde argumentiert, dass für \mathcal{C}_1 Netzwerke $2^{0.6\sqrt{N}}$ eine untere Schranke für die mittlere Anzahl der Attraktoren darstellt. Die Simulationen in Abb. 6.7 zeigen, dass allgemein 2^{aN^b} mit zwei Konstanten a und b ein gutes Fit darstellt.

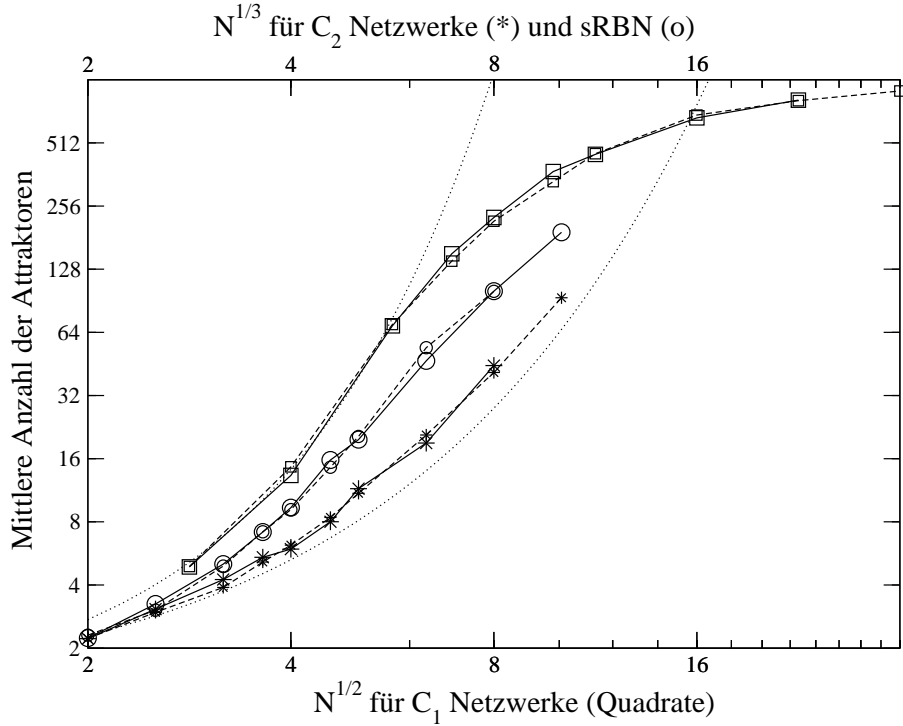


Abb. 6.7.: Mittlere Anzahl der Attraktoren aus der Simulation von jeweils 1000 Anfangszuständen in 1000 Netzwerken (gestrichelte Linien) und 2000 Netzwerken (durchgezogene Linien) für kanalisierende C_2 Netzwerke (*), sRBN (o) und C_1 Netzwerke (\square). Der Maßstab für die Kurven ist so gewählt, dass die Abszissen für die verschiedenen Kurven jeweils größenordnungsmäßig die Anzahl der relevanten Knoten angeben. Die punktierten Kurven entsprechen den Funktionen $2^{0.6 N_1^{0.5}}$ und $2^{0.55 N_1^{0.7}}$, wobei N_1 die Anzahl der Knoten in C_1 Netzwerken ist, das heißt, das Quadrat der Abszisse hier.

Auf der Grundlage der Computersimulationen von kleinen Netzwerken hatte Stuart Kauffman angenommen, dass die mittlere Anzahl der Attraktoren mit \sqrt{N} skaliert. Für kleine N stimmen die Daten in Abb. 6.7 mit dieser Annahme aus dem Jahre 1969 überein. Die analytischen Ergebnisse der letzten Jahre zeigen, dass die mittlere Anzahl der Attraktoren schneller als jedes Potenzgesetz mit der Netzwerkgröße ansteigt. Diese Erkenntnis wird durch Abb. 6.7 auch bestätigt. Die Ähnlichkeit in der Form der Kurven in der Abb. 6.7 ist wieder ein Indiz für die Richtigkeit der in dieser Arbeit entwickelten Vorstellung über die Struktur der RBN mit $K = 2$, über die relevanten Komponenten und die Implikationen für die dynamischen Eigenschaften.

7. Ausblick

Nachdem die Ergebnisse dieser Arbeit im Kapitel und am Anfang der jeweiligen Kapitel zusammengefasst waren und nachdem sinnvolle Erweiterungen des RBN - Modells im Kontext der Erforschung der genetischen Regulation im Kapitel 2 besprochen wurden, möchte ich an dieser Stelle einen Ausblick auf die mögliche Entwicklung des Forschungsgebietes niederschreiben.

Zunächst ist das RBN - Modell komplex und umfassend genug und gleichzeitig handlich genug um verschiedene Methoden aus der statistischen Physik und aus der Mathematik darauf anzuwenden. Obwohl die Dynamik des Modells lange Zeit nicht verstanden war, ist die Liste der inzwischen auf dieses Modell erfolgreich angewendeten analytischen Methoden relativ lang: das direkte probabilistische Studium der Attraktoren in [Samu 03, Dros 05a], die Strukturuntersuchungen mit Hilfe von stochastischen Prozessen in [Kauf 05b, Paul 06], “annealed approximation” und “generating functions” - Methode bei der Untersuchung des Skalenverhaltens der skalenfreien booleschen Netzwerke in [Deok 06] oder die probabilistische Betrachtung des breit anwendbaren Ansatzes der Störungsausbreitung in [Samu 06].

In meinen Augen haben die Modelle der booleschen Zufallsnetzwerke vieles (Existenz von Universalität und Skalenverhalten, Rolle der stochastischen Fluktuationen, Multistationarität) gemeinsam mit den in benachbarten wissenschaftlichen Gemeinden viel studierten Modellen (siehe [Dick 02, Hinr 00, Muti 04] für ein paar Reviews) der Phasenübergänge in absorbierende Zustände (viele Reaktions - Diffusions - Systeme gehören dazu), der selbstorganisierten Kritikalität und der Wachstumsvorgänge. Da es ferner Beispiele, siehe z. B. [Odor 04], der Abbildung von Reaktions - Diffusions - Systemen auf Wachstumsmodelle und dergleichen gibt, vermute ich, dass durch vergleichende Studien von verschiedenen Modellen der kritischen Phänomene im Nichtgleichgewicht das Werkzeugkasten für boolesche Zufallsnetzwerke erweitert werden kann. Allgemeiner geht es um die Weiterentwicklung der statistischen Nichtgleichgewichtsphysik.

Im Hinblick auf die realistischere Beschreibung der genetischen Regulation, solange man bei der diskreten booleschen Beschreibung bleibt, sprechen die experimentellen Daten für eine Einschränkung auf eine bestimmte Netzwerktopologie [Albe 05] und bestimmte Knotenfunktionen wie “unate” Funktionen in [Gref 06]. Diese und vergleichbare Einschränkungen enthalten Funktionen, die im Wesentlichen unter Komposition abgeschlossene Klassen boolescher Funktionen bilden. So oder auf eine andere Weise werden keine sehr langen Attraktoren zugelassen im Gegensatz zum Fall kanalisierender Funktionen im Abschn. 6.2.2. Des Weiteren müssen die realistischen “zuverlässigen” Attraktoren, wie beispielsweise in [Klem 05b, Klem 05a] studiert, resistent gegen kleine Verzögerungen der Eingänge und gegen andere störende

7. Ausblick

Faktoren sein. Die robuste Funktion ist im Allgemeinen in biologischen Systemen gegeben, auch wenn einzelne sich nicht heilende Mutationen z.B. in *Drosophila melanogaster* sowohl in lebenden Organismen nachgewiesen als auch künstlich erzeugt wurden. Die Aufgabe, kleine robuste (selektiv attraktivere) Netzwerkbausteine im (durch realistische Forderungen eingeschränkten) Netzwerkensemble zu identifizieren, ist dann übrigens vergleichsweise einfach [Klem 05b]. Bei der Untersuchung spezifischer genetischer Regulationen in spezifischen Organismen bedarf es schließlich keiner Physikerausbildung.

Die Fragestellungen, bei denen zunächst generische Modelle untersucht werden müssen, bevor (wenn überhaupt) realistische Vorgänge modelliert werden können, sind die Evolution von genetischen Netzwerken und die Berücksichtigung des Einflusses der Umgebung. Da beide Aufgaben grundsätzlich Modelle mit immer mehr Bestandteilen nach sich ziehen, gehören sie im Prinzip zum Untersuchungsgegenstand der allgemeinen Systemtheorie bzw. der Komplexitätstheorie. Die Physik und die Komplexitätstheorie müssen dabei voneinander profitieren.

Anhang

A. Entwicklung der Theorie der Komplexität

Die Annalen der Erforschung der Komplexität im weiteren Sinne zusammenzustellen wäre bestimmt ein sehr schwieriges Unterfangen für einen Geschichtsforscher. Ich verfolge in diesem Abschnitt eher das Ziel, kurz und subjektiv den Prozess der Entstehung der aktuellen interdisziplinären Forschungsschwerpunkte in der Physik zu beleuchten, die unter dem Namen der Theorie der Komplexität zusammengefasst werden können. Es werden relevante wissenschaftliche Disziplinen, Namen, Begriffe und Daten genannt. Unter Berücksichtigung der ungeheuren Weite des Feldes ist es dabei nicht verwunderlich, dass Fachleute aus den vielen betroffenen wissenschaftlichen Disziplinen sich zum Teil unvereinbarer und widersprüchlicher Terminologie bedienen.

Zeitgenössische Philosophen fragen, woher die Komplexität kommt und wie sie sich den interessierten Fragen gegenüber verhält. Einige fragen, ob es überhaupt etwas Neues über die Komplexität zu erzählen gibt. Einige fragen, ob ein komplexes System durch Studieren zerstört würde. Einige fragen, ob die Komplexität für uns gut oder schlecht ist. Spinoza sagte einmal, dass der Akt des Verstehens, also eine sinnvolle subjektive Vereinfachung der Komplexität, ethisch (gut) sei. Wieder andere fragen, wo die Verantwortung der allgemeinen Systemtheorie und der Wissenschaft über die Komplexität bei der Begegnung der aktuellen Ereignisse, der Politik, der Wertminderung in der Gesellschaft etc. liegt.

Viele Fragen sind nicht neu. Nach der Lehre des griechischen Philosophen Heraklit entspringt alle Ordnung einem ewigen Widerstreit gegensätzlicher Kräfte [Pasl 92]. Dem griechischen Philosophen Demokrit, einem Vertreter und dem Begründer der Atomistik, kann die Idee zugeschrieben werden, dass die Atome mit ihren intrinsischen Formen und Geschwindigkeiten, wenn sie nur genügend viele sind und genügend viel Zeit haben, keine Intelligenz brauchen um unter anderem *zwangsläufig* sich zu Allem was man kennt zu organisieren. Später hat beispielsweise R. Descartes im fünften Kapitel von "Discours de la méthode" gezeigt bzw. spekuliert [Desc 85], dass wenn Gott genügend Chaos geschaffen hätte und seine Gesetze hätte anschließend wirken lassen, so würde dieses Chaos sich zur Materie der Erde und Planeten organisieren. (Wenngleich in der abendländischen Geistesgeschichte zur Erklärung der Ordnungsprozesse auf das Göttliche zurückgegriffen wird.) Adam Smith, der schottische Moralphilosoph und Begründer der klassischen Volkswirtschaftslehre spricht von der "unsichtbaren Hand", der *Selbstorganisation* nach modernem Verständnis, die beim Zusammenwirken der Menschen in Märkten das Gemeinwohl fördert, ohne dass jemand von ihnen das beabsichtigt hat. Im 19ten Jahrhundert stellt dann I.

Kant in “Allgemeine Naturgeschichte und Theorie des Himmels” die Frage, ob die Newton’sche Mechanik auch die Ordnung in der Natur erklären könne, er hat die Vorstellung, dass die Materie in sich ein Bestreben trägt “sich zu bilden” [Pasl 92]. Offenbar ist die Frage nach Mechanismen der Entstehung der Komplexität ziemlich alt.

Die Mathematik kennt seit dem 17ten Jahrhundert die Idee der wiederholten Selbstähnlichkeit, eine der vielen wegbereitenden Einsichten von Gottfried Leibniz. Als eine Krönung der Leibniz’schen Selbstähnlichkeit hat 1975 B. Mandelbrot den Begriff *Fraktal* geprägt [Mand 82]. Die Beschreibung von komplexen Systemen kann man sich heutzutage ohne Benutzung von zu Fraktalen verwandten Konzepten wie Skalierung kaum vorstellen. Neben der fraktalen Geometrie sind auch andere neuere Teilgebiete der Mathematik wie die Wahrscheinlichkeitstheorie, die Chaostheorie und die numerische Mathematik aus der Untersuchung der Komplexität nicht wegzudenken. Die eigene Entwicklung dieser mathematischen Disziplinen steht im engen Zusammenhang mit der Untersuchung der Komplexität. Ohne zu übertreiben, die Fragestellungen im Zusammenhang mit der Komplexität bemühen die meisten mathematischen Disziplinen.

Der englische Philosoph G. H. Lewes hat 1875 in Verbindung mit der Beschreibung von Produkten bestimmter chemischer Reaktionen, als einer der ersten, vom *emergenten* Verhalten gesprochen [Lewe 75], wenn qualitativ neue Eigenschaften dieser Produkte kaum als eine Überlagerung der bekannten Eigenschaften der Systemkomponenten verstanden werden konnten. Der deutsche Philosoph N. Hartmann hat im Kontext seiner Gesetze über die Kategorien der Komplexität [Hart 40, Steg 69] das moderne Verständnis der Emergenz, des kategorischen Novums, eingeführt; seine Arbeiten¹ haben den Vater der allgemeinen Systemtheorie Ludwig von Bertalanffy stark beeinflusst.

Als in den 20er Jahren des 20ten Jahrhunderts emergenter Evolutionismus (z. B. von C. Lloyd Morgan, 1923 [Smut 26]) dem Reduktionismus² entgegengestellt und mit mäßigem Zugewinn an Einsichten polemisiert wurde [Gold 99b], hat der Schotte D’Arcy Thompson 1917 ein einzigartig weitsichtiges Buch mit dem Titel “Über Wachstum und Form” veröffentlicht [Thom 17]. In diesem Buch beschrieb er mit damaligen mathematischen Mitteln biologische Formen, von der Breitenverteilung der Zebrastreifen über Allgegenwärtigkeit der Spirale wie auf Seemuscheln bis zu geometrischen Ähnlichkeiten der Hautmuster auf verwandten Fischarten. Auf Seite 270 sagt er: “we rise from the conception of form to an understanding of the forces

¹Einer der letzten Universalgebildeten (polymath) Nicolai Hartmann hat die Aufgabe der Philosophie in der Entdeckung der Strukturgesetze der realen Welt (und keiner Scheinwelt) gesehen. Er hat auf die Möglichkeit eines partiellen Begreifens hingewiesen, wobei eine unendliche Irrationalität zurückbleibt.

²Zurückführbarkeit von (allen) Phänomenen auf eine “grundlegendste” Wissenschaft, die diese Phänomene dann im Prinzip verständlich macht.

which gave rise to it”³. Diese Aussage entspricht dem *Sinn der Fragestellungen der modernen Wissenschaft über die Komplexität*.

Die physikalische Untersuchung der Komplexität im engeren Sinne muss im Kontext der Entwicklung der Weltanschauung und aus der historischen Perspektive im Zusammenhang mit der Entwicklung der Allgemeinen Systemtheorie betrachtet werden.

Vermutlich im sechsten Jahrhundert vor Christi hat Laozi im daoistischen Urwerk “Tao Te King” “*systematische*” Weisheiten gelehrt, die im ursprünglichen Kontext beeindruckender sind: “... Darum führt die Richtung auf das Nichtsein zum Schauen des wunderbaren Wesens. ... Der Sinn richtet sich nach sich selber...” oder “ohne die einzelnen Bestandteile eines Wagens gibt es keinen Wagen” [Laoz 06a] und “Der Sinn erzeugt die Eins (Liebe). Die Eins erzeugt die Zwei (Genügsamkeit). Die Zwei erzeugt die Drei (Zurückhaltung). Die Drei erzeugt zehntausend Dinge. Die zehntausend Dinge beinhalten den Schatten und verbreiten das Licht, und die beiden Kräfte zusammen (die verflochtenen “yin” und “yang”) strömen in Harmonie” [Laoz 06b]. Nennenswert ist auch die von G. Leibniz, 400 Jahre nach ähnlichen Gedanken vom spanischen Scholastiker Raymundus Lullus (1235 - 1315), angestrebte universelle Begriffsbildung “*characteristica universalis*”.

Am Anfang des 20ten Jahrhunderts war Arzt, Wissenschaftler und Revolutionär A. Bogdanov — sein Werk “Tektologie”⁴ in drei Bänden erschien 1922 — seiner Zeit um 30 Jahre voraus [Dudl 96, Bigg 98], als er vorschlug, die Sozialwissenschaften und die Naturwissenschaften mit ihren unvereinbaren Terminologien dadurch unter einen Hut zu bringen, dass die Wissenschaften als ein System von Relationen gesehen werden sollten, die einheitlich nach denselben Prinzipien organisiert sind. Etwa zur gleichen Zeit wurde die Quantenmechanik entwickelt, die nach einer Kapitelüberschrift des Buches “Science and Humanism” von E. Schrödinger [Schr 52] mit “form, not substance - the fundamental concept”, also mit Beziehungen statt mit absoluten Einheiten arbeitet. Auch zu dieser Zeit fand der Begriff “*Holismus*” seinen Platz im Lexikon. Nach der entsprechenden Sichtweise ist das Ganze mehr als die Summe seiner Einzelteile. Es gab auch viele andere Anzeichen für die heranreifende Bestrebung nach einer teilweise möglichen Zusammenführung aller Wissenschaften wie der 1935 an der Sorbonne in Paris stattgefundene internationale Kongress “Einheit der Wissenschaften”⁵. Der beginnende, zum Teil wieder entdeckte, *Paradigmenwechsel* bestand nach Fritjof Capra [Capr 00] aus fünf Neuerungen und sollte nicht nur die Wissenschaft betreffen: das Ganze statt der Bestandteile zu betrachten, statt Strukturen erzeugende Prozesse zu untersuchen, relative (blickwinkelabhängige, duale, erkenntnistheoretische) statt objektiver Erkenntnisse, ver-

³Wir bewegen uns von einer Beschreibung von Formen zu einem Verständnis der Kräfte, die sie erscheinen ließen.

⁴Der Begriff Tektologie (tectology) wurde von A. Bogdanov eingeführt und stammt vom griechischen “tekein” (zusammenbauen) bzw. später lateinischem “tectum” (Überdachung). In der westlichen Welt kaum verwendet, bis auf eine ähnliche Bedeutung in der Morphologie von Haeckel.

⁵Vgl. <http://www.muellerscience.com/spezialitaeten/System/Systemwissenschaft.htm>.

flochtene (netzwerkartige, gegenseitige, nicht lineare) Beziehungen statt linearen anhäufenden Aufbaus der Systeme und Erkenntnisse, annähernde (metaphorische) statt exakter (wahrer) Beschreibung.

In Worten von Russell L. Ackoff, einem Philosophen und Denker in der mathematischen Entscheidungsvorbereitung (operations research), der manchmal wegen seiner wegbereitenden Arbeiten auch als Dekan der systematischen Denkweise bezeichnet wird, „ist *systematische Denkweise* ein Zusammenschluss der Analyse (auseinander nehmen) und der Synthese (die Rolle der Einzelteile im größeren System verstehen), abhängig davon, ob unser Ziel (Struktur-) Wissen oder Verstehen ist“⁶. In Worten vom Physiker David Bohm, der 1952 sein „holistisches“ Quantenpotential verteidigt: Das Quantenpotential steuert die Teilchenbewegung, indem es „aktive Information über die gesamte Umgebung liefert, wie Radarsignale“, die die Motoren eines Schiffes „steuern“ [Bohm 87]. Und Albert Einstein sagte einmal: „A human being is part of the Whole. . . He experiences himself. . .“⁷ [Eins 54].

Zwei einflussreiche Arbeiten haben maßgeblich zu der modernen allgemeinen Systemlehre beigetragen. 1948 hat Norbert Wiener, ein begnadeter Mathematiker aus Amerika, ein Pionier in der Untersuchung der in dieser Arbeit relevanten stochastischen Prozesse wie die Brown'sche Bewegung, das Wissen über regulatorische und nachrichtentheoretische Prozesse zusammengefasst und die Kybernetik⁸ begründet. Die Kybernetik betrachtet sich selbst regulierende Systeme, die aus ihren Teilen und ihrer Organisation bestehen. Getreu dem Buchtitel „Cybernetics or Control and Communication in the Animal and the Machine“ [Wien 48] beschreibt der Autor N. Wiener einheitlich die regulatorischen Kreise (negative, positive, corrective feedback), die Irreversibilität, die Rolle der Information — in elektrischen Kreisen, im menschlichen Körper, in der Gesellschaft. Er betont funktionale zielgerichtete Beziehungen, die die Essenz der Neuerungen ausmachen. Eine bekannte Umschrei-

⁶Siehe http://ackoffcenter.blogs.com/ackoff_center_weblog/2005/06/from_mechanisti.html zum Vortrag von R. Ackoff auf der Konferenz „Systems Thinking in Action Conference“ in 1993. R. Ackoff sieht die systematische Denkweise und die damit verbundene Informationsgesellschaft als die erste große gesellschaftliche Revolution nach der industriellen Revolution. Er propagiert die Ansichten seines amerikanischen Lehrers, des Philosophen Arthur Singer, Jr., der 1898 (ignoriert von der etablierten Wissenschaft bis zum Erscheinen der allgemeinen Systemtheorie) eingesehen und konsequent weiter verfolgt hat, dass „aus einer Eichel eine Eiche nur in einer entsprechenden Umgebung wachsen kann“. Also ist die Eichel nicht die „Ursache“ der Eiche, die Umgebung ist nicht zu vernachlässigen, man kann diesen Prozess auf viele komplementäre Weisen beschreiben und man kommt zu einer zweckgebundenen Beschreibung und einer teleologischen Weltanschauung. Unter diesem Gesichtspunkt werden Methoden den Theorien vorgezogen, was immer noch kontrovers ist. R. Ackoff spricht in seinem Vortrag auch von einem Buch, das in den 1960ern der Bestseller der größten universitären Buchhandlung der Welt, derjenigen an der Universität von Kalifornien, war. Nicht die Bibel. Kein Wörterbuch. Keine Straßenkarte. „I Ching“ über den Zen Buddhismus, dessen Philosophie der Theorie über die Komplexität sehr ähnlich ist. Sie beruht auf einer dynamischen Balance zwischen gegensätzlichen Kräften, auf einem Begreifen der Ereignisse als Prozesse, auf einem Akzeptieren der unvermeidlichen Änderungen.

⁷Ein Mensch ist ein Teil des Ganzen. . . Er erlebt sich selbst. . .“.

⁸Das Wort Kybernetik stammt vom griechischen „kybernêtes“ (Steuermann).

bung der Kybernetik von einem ihrer Pioniere Louis Couffignal lautet “the art of assuring efficiency of action”⁹. Die mathematischen Methoden der Kybernetik basieren auf Differentialgleichungen für zeitliche Veränderungen des Systemzustands.

Ein anderer Kybernetiker und Erforscher der komplexen Systeme, der englische Psychologe William R. Ashby, hatte einen großen Einfluss auf N. Wiener, L. von Bertalanffy, S. Kauffman und andere und ist auch dafür bekannt, dass in seinen Veröffentlichungen der Begriff *Selbstorganisation* zum ersten Mal in Druckform vorkommt [Ashb 47].

Die Systemwissenschaft könnte man mit einiger Vorsicht auf die gleiche umfassende Stufe stellen wie die Philosophie und die Mathematik, und diese Tendenz nimmt langsam zu. Die Literaturliste zur allgemeinen Systemtheorie, meist auf Englisch, umfasst mehrere Hundert Titel¹⁰.

Die einflussreichsten Arbeiten vom allgemein als Vater der allgemeinen Systemtheorie (GST) [Bert 45, Bert 54] angesehenen österreichischen Biologen K. Ludwig von Bertalanffy kamen Anfang der 50er Jahre zur richtigen Zeit, denn “etwas lag in der Luft”. (Er hat einige seiner Ideen bereits in den 30er Jahren geäußert.) Die Grundprinzipien waren in der Biologie erprobt und umfassten eine ganzheitliche Betrachtung auf allen Organisationsstufen, die Prinzipien “dynamisch statt statisch” und “aktiv statt reaktiv” im Bezug auf jeden Organismus in seiner Umgebung. Die allgemeine Systemtheorie sollte ein logisch - mathematisches Gebiet werden, das übergreifende komplementäre Charakteristiken und Sichtweisen studiert und logische Homologien in allen Wissenschaften (z. B. hierarchische Ordnung) bedingt. Der englische Wirtschaftswissenschaftler K. E. Boulding hat das Ziel der GST als verallgemeinerte Ohren, die die spezialisierte Taubheit überwinden sollten, umschrieben. Er behauptete, dass es für jeden Zweck und jede Abstraktionsebene einen optimalen Grad der Allgemeingültigkeit gibt, irgendwo zwischen dem sinnlos Offensichtlichen und dem gegenstandslos Allgemeinen. Von Bertalanffy, Boulding, Gerard und Rapaport haben 1954 “Society for the Advancement of General Systems Theory” gegründet, die heutzutage “International Society for the Systems Sciences” heißt und ein Internetportal unter <http://www.iss.org/> pflegt¹¹. Von Bertalanffy hatte eine Vision, die über die Effizienz der Wissenskommunikation hinausging, er verfolgte eine humanistische Perspektive, in der ein Individuum harmonisch in seiner Umgebung existiert.

Laut Bela H. Banathy ist die systematische Weltanschauung kein abstraktes Gebilde und kann konsequent verfolgt werden. Diese Anregung haben sich viele Wis-

⁹Die Kunst, effizient zu beeinflussen.

¹⁰Vergleiche <http://www.muellerscience.com/spezialitaeten/System>, dort die Datei *Systeme_Geschichte_Uebersicht.htm*. Eine sehr interessante Zusammenstellung von Dr. phil. Roland Müller. Es wird schnell klar, dass es unmöglich ist, alle Vordenker beim Namen zu nennen.

¹¹Eine nette Zusammenstellung der einflussreichen Systemtheoretiker im weiteren Sinne findet man im Jahr 2006 unter <http://pespmc1.vub.ac.be/CSTHINK.html>, unter http://business.nmsu.edu/~dboje/655/655_overview.htm gibt es eine Literaturzusammenstellung von Ph. D. David Boje, die insbesondere auf verschiedene Schulen innerhalb der Systemtheorie eingeht.

senschaftler zum Teil unabhängig von den Arbeiten von Wiener und von Bertalanffy auf die Fahnen geschrieben. Folgende unvollständige subjektive Auswahl der wissenschaftlichen Disziplinen belegt das, bevor auf die im engeren Sinn Komplexitätstheorie - relevanten Gebiete näher eingegangen wird.¹²

In der ersten Hälfte des 19ten Jahrhunderts erfahren die Wahrscheinlichkeitstheorie, Statistik, Mengenlehre, Topologie usw. ungeahnte Aufschwünge. Die moderne Wahrscheinlichkeitstheorie erschien 1929 aus der Feder des sowjetischen Mathematikers Andrey N. Kolmogorov, der später auch große Beiträge zur Informationstheorie, zur Turbulenz u. v. m. leistete¹³. Dass prinzipiell nicht alle mathematischen Fragen anhand von Axiomen beantwortet werden können, hat der österreichische Mathematiker Kurt Gödel 1931 gezeigt. Die heute wenig bekannte Allgemeine Semantik wurde 1933 vom polnischen Denker Alfred Korzybski auf 800 Seiten entwickelt [Korz 95]. Sie sucht nach praktischen Kriterien der Vernunft in der sprachlichen Kommunikation und enthält systemische Konzepte wie die Notwendigkeit, sich der gemachten Abstraktionen bewusst zu sein, wie "Die Karte ist nicht gleich der Landschaft" ("the map is not the territory") und wie die Existenz intrinsischer Grenzen der Wissenskommunikation ("time binding"). Einer der Väter der modernen Informatik, der englische Mathematiker Alan Turing, hat 1936 die Turing Maschine beschrieben [Turi 36]. Die Geschichte der neuronalen Netzwerke beginnt etwa 1940 mit Arbeiten vom Neurophysiologen Warren S. McCulloch und vom Logiker Walter Pitts [McCu 43]. Anfang der 40er Jahre wird ein selbstreproduzierender zellulärer Automat vom exzentrischen universal gebildeten ungarischen Mathematiker John von Neumann beschrieben, was den Anfang der Untersuchung der zellulären Automaten darstellt. Von Neumann ist auch am Bau der ersten Computer beteiligt und ist 1944 der Begründer der Spieltheorie [Morg 47]. Eine Analyse der funktionellen Wechselbeziehungen in der Wirtschaft, mathematisch und mit ei-

¹² Eine Fülle relevanter Begriffe dokumentiert die Richtung der Forschung: systemische Ontologie, Epistemologie und Ethik, allgemeine Semantik, Konstruktivismus, kontrollierte Benutzung der Metapher, Analogien und Isomorphismen, interdisziplinäre Aspekte der Komplexität, adaptive Systeme, allgemeine Taxonomie der Systeme, Adaptivität, Aura, Autogenese, Autonomie, Autopoiesis, Beobachter, Bionik, Chaos, critical systems thinking, Dissipation, dynamische Systeme, Emergenz, Emergie, Entscheidungsfindungssystem, Evolution, Feedback, Fließgleichgewicht, ganzheitlich, Hierarchie, Heterarchie, Homeostase, Homeostat, homöostatisches Gleichgewicht, Hyperzyklus, Implosion, Information, Informationstheorie, Irreversibilität, komplexe Systeme, Komplexitätstheorie, Kritikalität, künstliche Intelligenz, Kodierung, Kommunikation, Kontrollsysteme, Kooperation, lernende Systeme, Markov'sche Prozesse und Matrizen, mathematische Entscheidungsvorbereitung, mathematische Systemtheorie, Metamodellierung, Metastabilität, Metasystem, Morphogenese, Motive, Muster, neuronale, skalenfreie usw. Netzwerke, neuronale Organisation, Nichtlinearität, operationelle Abgeschlossenheit, offene Systeme, Ordnungsparameter, Perkolation, Rand- und Nebenbedingungen, Regelung, Robotik, Selbstorganisation, Selbstreferenz, Selektion, Simultaneität, Skalierung, Stabilität, stochastische Prozesse, Simulation, Stress, Synergie, Systementwurf, Systemmanagement, Transformität, Unvollständigkeit, Umgebung, Vielfältigkeit, Wahrscheinlichkeit, Zufall und so weiter. Vergleiche <http://projects.issis.org>.

¹³Vergleiche <http://homepages.cwi.nl/~paulv/kolmogorov.biography.html>,
<http://www.kolmogorov.com/>.

nem der ersten Computer, hat Wassily W. Leontief etwa 1941 eingeführt [Leon 51]. Herbert A. Simon mit seiner wirtschaftlichen Entscheidungstheorie (input - output - method) (1943) und späteren Beiträgen zum rationalen Aufbau der Organisationen war ein anderer einflussreicher amerikanischer Forscher, der vom systematischen Ansatz Gebrauch gemacht hat. Der Begriff Futurologie bzw. Zukunftsforschung als systematische Behandlung von Zukunftsfragen wurde 1943 von Ossip K. Flechtheim in den USA geprägt¹⁴. Die Kybernetik entwickelte sich 1945 - 1947 (W. R. Ashby, N. Wiener). Die Thermodynamik erfährt zahlreiche Ausweitungen in Richtung offener Systeme und irreversibler Prozesse. Die irreversible Thermodynamik wurde 1947 in der Dissertation vom belgischen Physiker und Chemiker Ilya Prigogine begründet, ab etwa 1968 wird er sich später mit der Selbstorganisation und den von ihm geprägten dissipativen Strukturen beschäftigen. Der Informationsbegriff wird eingeführt. Das Fundament der Informationstheorie wurde 1948 von Claude E. Shannon gebaut [Shan 48] und von Warren Weaver popularisiert. Das Buch mit dem Titel "Systemtheorie der elektrischen Nachrichtenübertragung" von Karl Küpfmüller ist 1949 erschienen [Küpf 49]. Das Studium des sozialen Verhaltens von Termiten führte den französischen Biologen Pierre - Paul Grassé 1959 zur Formulierung der Stigmergie, einer Disziplin, die Systeme studiert, deren Teile durch Änderungen ihrer lokalen Umgebung miteinander kommunizieren [Gras 59]. Eng verwandt mit der Kybernetik ist die angewandte Kontrolltheorie aus dem Ingenieurwesen und der Mathematik, die sich als systematische Disziplin Ende der 50er Jahre etabliert hat, sich mit stabiler Regulation auf Sollwerte beschäftigt und moderne mathematische Konzepte wie fuzzy - Logik (Prof. Lotfi Zadeh, 1965) ausnutzt.

Die nächste große Anstrengung der wissenschaftlichen Gemeinde fing in den 70er bzw. 80er Jahren an, als die Komplexität mit modernen Mitteln an der Wurzel gepackt wurde. Man kann grob zwei komplementäre Strömungen unterscheiden. Die bekanntere sucht nach in einem gewissen Sinne universellen Gesetzmäßigkeiten der Komplexität und ihrer selbstorganisierten Entstehung und betont die involvierten Prozesse. Den Grundstein hat hier Stuart Kauffman mit seiner Idee "Ordnung gratis" gelegt. Die andere Strömung heißt Synergetik¹⁵. Die Wissenschaft Synergetik wird seit den Arbeiten vom deutschen Physiker H. Haken in den 70er Jahren [Hake 83] auch als ein alternativer Name zur Theorie der Komplexität verstanden, mit einer Betonung auf nicht lineare Beziehungen der Systemelemente untereinander¹⁶. Die berühmtesten Verfechter der Hypothese, dass der kooperative Nutzen (excess stemming from cooperation) die Komplexität bedingt, sind der Amerikaner Peter A. Corning und später der englische Biologe John Maynard Smith und der

¹⁴Vergleiche <http://www.izt.de/> und insbesondere http://www.izt.de/pdfs/IZT_AB_23.pdf.

¹⁵Das griechische Wort "synergia" bedeutet "gemeinsames Werk". Der Begriff Synergieeffekt wurde vermutlich vom amerikanischen Architekten und Philosophen Buckminster Fuller geprägt. Er propagierte globale Sichtweise und wurde bekannt durch seine geodätischen Kuppeln.

¹⁶Haken hat sich ursprünglich mit der statistischen Physik der Nichtgleichgewichtssysteme beschäftigt, insbesondere mit dem Laser. Vergleiche dazu den Text über I. Prigogine. Dabei konnten die Paradigmen wie Ordnungsparameter, Versklavungsprinzip, Phasenübergänge, Selbstorganisation vorangetrieben werden.

ungarische Biochemiker E. Szathmáry [Corn 83, Smit 95]. Sie haben die Idee der “Gruppenselektion” in der Evolution untersucht und verallgemeinert. Sie betonen kumulative, zielgerichtete und (historisch) umgebungsbedingte Aspekte der Komplexität. Die kooperierenden Subjekte an sich sind die Ursache des Trends zum komplexen Verhalten. Diese Synergien bestehen aus den Komponenten Materie, Energie, Information, die man auch als Elemente, Beziehungen und Ziele (das Ganze) (aus der minimalistischen Definition der GST von Bertalanffys) umschreiben kann. Einige Ansätze des viel umfassenden synergetischen Konzepts sind: historische Betrachtung verdient mehr Beachtung (Zusammenspiel des Zufalls und der Notwendigkeit (“chance and necessity”)), Synergien scheinen während der Evolution zwangsläufig zu entstehen (beispielsweise sind Eukaryoten etwa 15-fach effizienter in der Energieumsetzung als die Prokaryoten), die Kausalkette kann induktiv und deduktiv untersucht werden (“upward and downward causation”).

Eine sinnvolle axiomatische Definition dessen, was die Theorie der Komplexität ausmacht, ist nicht möglich. Es ist jedoch möglich, grobe Kriterien aufzustellen, nach denen ein Forschungsthema zur Theorie der Komplexität zu zählen ist. Aus einer historischen bzw. einfach einer umfassenderen Perspektive werde ich bis zum Ende dieses Anhangs auch chaotische Systeme als komplex ansehen. Für die angesprochene Zuordnung reicht es aus, wenn eines der folgenden Kriterien erfüllt ist.

- Reichhaltige Verhaltensmuster im untersuchten System. Folgende Eigenschaften fallen in diese Kategorie:
 - Viele Systemkomponenten: Diese Bedingung betrifft wahlweise auch die Wechselwirkungen im System und ist offensichtlich.
 - Offenheit des Systems: Systeme fern vom Gleichgewicht (Systeme mit Dissipation oder Materie - Fluss oder Energiefluss durch das System) sind häufig komplex.
 - Nichtlinearität bzw. verflochtene Wechselwirkungen im System: dann können sich lokale Prozesse global auswirken, z. B. durch exponentielles Auseinanderlaufen der benachbarten Trajektorien im Phasenraum.
 - Nichtlokalität (zeitlich und/oder räumlich): Systeme mit Gedächtnis, mit pfadabhängigen Übergängen im Phasenraum, mit nicht - Markov’schen Prozessen neigen zum komplexen Verhalten.
 - Indeterminismus: stochastische Prozesse, asynchrone Wechselwirkungen, zufällige Prozesse, fuzzy - Logik etc. werden damit gemeint.
 - Viele Schichten: In den meisten komplexen Systemen kann man das lokale und das globale Verhalten unterscheiden; andere mögliche Stufen der Wechselbeziehungen begünstigen die Komplexität.
 - Alles andere, was Komplexität fördert.

- Geordnete bzw. strukturierbare Verhaltensmuster. Falls die Ordnung nicht eindeutig definierbar ist, wird hier explizit auf die Existenz eines Beobachters verwiesen. In dieser Kategorie findet man:
 - Attraktoren: Die Möglichkeit für ein Zusammenziehen des Phasenraumvolumens im Laufe der Dynamik lässt eine Strukturierung des Phasenraums zu.
 - Selbstorganisation, Emergenz: In einem bestimmten Sinne kann die Ordnung im System zunehmen.
 - Selbstregulation, adaptives Verhalten, Stabilität, Robustheit: In einem gewissen Sinne passt sich das System der Umgebung an bzw. die innere Struktur kann erhalten oder verteidigt werden. Homöostase (Erhalten der Funktion fern vom Gleichgewicht), Metastabilität, Fließgleichgewicht sind möglich.
 - Hierarchien, Skalierung: Im Gegensatz zu “vielen Schichten” oben wird hier eine durch die Dynamik des Systems bedingte (emergente) Identifizierbarkeit und einfache Beschreibbarkeit einer Vergrößerung (eines Herauszoomens) im System gemeint.

Nun komme ich zu den Theorien, die in das beschriebene Schema der Komplexitätstheorie passen und physikalisch interessant sind. Letzteres bedeutet, dass auf grundsätzlich anwendungsorientierte Theorien u. Ä. wie die Expertensysteme in der Wirtschaft oder Konnektionismus im Bereich der künstlichen Intelligenz, Autopoiesis usw. nicht eingegangen wird. Auch nicht auf verwandte “metaphysische” Themen wie die Kybernetik zweiter Art von H. von Foerster (die den Beobachter zum System dazuzählt und mit Beobachtungen operiert [Von 03]) oder die so genannte “unaussprechliche Ethik”.

Spontane Strukturbildung in (chemischen) Reaktions - Diffusions - Systemen wurde erstmals 1952 von A. Turing beschrieben [Turi 52]. Die Perkolationstheorie, die, in Kürze, verbundene Cluster auf Zufallsgraphen studiert und Phasenübergänge¹⁷ beschreibt, wurde 1957 von den Mathematikern S. Broadbent und John M. Hammersley begründet [Broa 57]. Viele Ergebnisse und Anwendungen folgten wegen des Schwierigkeitsgrades der Probleme allerdings erst in den 70er und 80er Jahren [Stau 94]. Trotz der ursprünglich geometrischen Natur der Probleme in der Perkolationstheorie können damit¹⁸ einheitlich zahlreiche kritische Prozesse in kom-

¹⁷An der kritischen Schwelle des Phasenübergangs entstehen verbundene Cluster von der Systemgröße.

¹⁸Im Hinblick auf Anwendungen sind in der Perkolationstheorie Ergebnisse zur so genannten Universalitätsklasse der gerichteten Perkolation besonders wichtig. Gerichtete Perkolation (directed percolation, DP) bezieht sich auf anisotrope Verbindungen der Knoten. Sie spielt für die Nichtgleichgewichtsphänomene eine Rolle, die mit der Rolle des Ising - Modells in der Gleichgewichtsphysik vergleichbar ist. Allerdings sind die kritischen Exponenten der DP nicht einmal in einer Dimension analytisch bekannt. Außerdem bleiben die quantitativen experimentellen Bestätigungen für die Existenz des reinen DP - Verhaltens weitgehend aus.

plexen Systemen beschrieben werden, von der Epidemien - Ausbreitung (Phasenübergang Eindämmung - Epidemie) bis zu Quarks - Confinement in der Kernphysik; neuere Beispiele stellen die weiter unten erwähnten Modelle (bestimmte zelluläre Automaten, CMLs, Kontaktprozesse, Wachstumsprozesse, Waldbrandmodelle etc.) dar, siehe Seite 49 in [Hinr 00] für Referenzen.

Der französische Mathematiker René Thom hat in den 60er Jahren die Katastrophentheorie entwickelt, die vom geborenen japanisch - britischen Mathematiker Sir Erik Christopher Zeeman in den 70ern popularisiert wurde. Diese Theorie betrachtet "katastrophale" Änderungen der Systemeigenschaften und ist ein Teil der Singularitätstheorie. Sie macht Gebrauch von der modernen Differentialtopologie, der algebraischen Geometrie und der Theorie dynamischer Systeme. Die Katastrophentheorie hat zur Klärung vieler Probleme in der Quantenoptik, der Thermodynamik und der Streutheorie beigetragen. Insbesondere sind in der Natur zahlreiche Beispiele für die so genannte cusp - Katastrophe und die Tangenten - Katastrophe bekannt.

Die deterministische Chaostheorie in niedrig - dimensional Systemen ist erst im Computerzeitalter auf allgemeines Interesse gestoßen¹⁹. Das klassische Beispiel ist das vom Meteorologen E. N. Lorenz 1963 studierte barotropische Modell der Atmosphäre, das hoch empfindlich auf Anfangsbedingungen reagiert (seltsamer Lorenz-Attraktor). Dieses Phänomen ist unter dem Namen Schmetterling - Effekt bekannt. Im allgemeineren Fall der räumlich verteilten Freiheitsgrade in komplexen chaotischen Systemen spricht man vom räumlich - zeitlichen Chaos, das auch Reaktions - Diffusions - Systeme zeigen. Klassische Beispiele sind das chaotische Regime der Belousov - Zhabotinsky Reaktion und die Taylor'schen Wirbel in der Geophysik [Cros 93].

Die Zufallsgraphen, die für die vorliegende Arbeit relevant sind (sie gehören zu extremalen Problemen bzw. sie besitzen Schwelleneigenschaften bzw. sie zeigen physikalisch gesehen einen Phasenübergang), wurden 1960 vom legendären ungarischen Mathematiker Paul Erdős und dem anderen ungarischen Mathematiker A. Rényi eingeführt. Die Renormierungsmethoden und Skalierungsansätze standen den Wissenschaftlern nach ihrer Entwicklung in den 50er und 60er Jahren, spätestens jedoch seit den Arbeiten des amerikanischen Physikers Kenneth G. Wilson über die Phasenübergänge zweiter Ordnung in den 70er Jahren zur Verfügung²⁰.

¹⁹Wesentliche Beiträge wurden bereits von J. Hadamard (Hadamard's billiards, 1898), A. Lyapunov (Stabilität, 1884), H. Poincaré (Dreikörper-Problem, 1887), J. D. Birkhoff (Ergodizitätstheorie, 1931), Mary Cartwright und J.E. Littlewood (ein chaotisches System, 1938), A. N. Kolmogorov (zur Turbulenz, 1941; KAM - Theorem, 1953), S. Smale (horseshoe - Abbildung, 1967), D.P. Ruelle (prägt den Begriff "seltsame Attraktoren", zur Turbulenz, 1971), J. A. Yorke (prägt den Begriff "Chaos", 1975) geleistet [Ott 02]. Ein chaotisches dynamisches System muss nach Definition sensitiv zu Anfangsbedingungen, im Phasenraum dicht liegende periodische Trajektorien aufweisend und topologisch mischend sein.

²⁰Falls man eine Liste der Werkzeuge, die bei der Untersuchung der Komplexität zum Einsatz kommen erstellen müsste, dann wären darin die meisten in der Philosophie, der Mathematik, der Physik, der Theorie der Komplexität usw. entwickelten Methoden enthalten. Einige möchte ich in Form von Stichworten benennen. Die Liste ist zusammengestellt mit Hilfe von

Die meisten komplexen Systeme sind Nichtgleichgewichtssysteme. In Worten vom polnischen Mathematiker Stanislaw Ulam wäre ihre Theorie die Theorie der Nichtelephanten. Die Nichtgleichgewichtssysteme wurden zunächst in den 70er Jahren studiert um Relaxation ins Gleichgewicht zu beschreiben [Hohe 77]. Später verlagerte sich das Interesse auf offene Systeme mit Energie- oder Teilchenfluss durch das System, Phasenübergänge im Nichtgleichgewicht und deren Universalitätsklassen, Skalenverhalten mit zusätzlichen im Vergleich zum Gleichgewicht kritischen Exponenten, Irreversibilität durch absorbierende Zustände [Odor 04]²¹. In den 70ern wird wie bereits erwähnt die Theorie der Nichtgleichgewichtsthermodynamik von I. Prigogine entwickelt und der Begriff dissipativer Strukturen geprägt, die mittels bestimmter (irreversibler) Energieumsetzung ihre innere Ordnung erhalten, die ansonsten zerfällt. Die dissipativen Strukturen sind ein anderer Ausdruck für die Selbstorganisation²². Die Theorie der Hierarchie ist ein Dialekt der allgemeinen Systemtheorie und wurde durch Herbert Simon, Ilya Prigogine und den Psychologen Jean Piaget entwickelt [Ahl 96, Simo 62]. Sie findet Anwendung seit den 70er Jahren z. B. in der Ökologie, Psychologie und basiert auf Konzepten wie definitionsabhängige Hierarchiestufen (ihre Struktur definiert die Komplexität), Skalierung, Beobachter gegen Beobachtung, Prozess gegen Struktur, das Element gegen das Ganze (beides zusammen auch Holon genannt). Höhere Hierarchiestufen bedeuten Zwangsbedingungen wie die bedingte Offenheit des Systems für die darunter liegenden Hierarchiestufen.

Eine ganzheitliche formalisierte Beschreibung der lebenden Systeme mit Hilfe der symbolischen, an die elektrischen Schaltkreise angelehnten, Energie - System

<http://www.necsi.org/guide/points.html>, wobei ich, wie in der Einleitung zu dieser Arbeit beschrieben, die Komplexität etwas anders verstehe. Die Stichworte sind: holistische Betrachtung, Skalierungsgesetze, kritische Phänomene, Nichtgleichgewicht, Differentialgleichungen, (feldtheoretische) Renormierungsmethoden, generierende Funktion, Transferoperator, Perkolation, Fraktale, Bäume, Netzwerke, Graphen, Gitter, Hierarchie, Topologie, Clustering, Wechselspiel der Topologie und Funktion, Abbildungen auf bereits studierte Probleme, Chaostheorie, Ansatz der dynamischen Antwort, Modelle mit nichtlinearer Rückkoppelung, Reaktionen - Diffusions - Modelle, lokale Aktivierung - langreichweitige Inhibition - Modelle, neuronale Netzwerke, Boole'sche Netzwerke, Kohonen - Netzwerke, Spingläser, Stabilitätsanalyse, Bifurkationen, Attraktoren, algorithmische Komplexität, Entropie, Information, numerische Simulationen, Monte Carlo, Quenching und Annealing, zelluläre Automaten, genetische Algorithmen, evolutionäre Algorithmen, Wettbewerb, Frustration, Fitness, stochastische Prozesse, Statistik, Kombinatorik, Intuition.

²¹Phasenübergänge im Nichtgleichgewicht kommen z. B. vor in Modellen der Populationsdynamik (Albano, 1994), der Epidemien - Ausbreitung (Ligget, 1985; Mollison, 1977), der Katalyse (Ziff et al., 1986), in Straßenverkehrsmodellen (Chowdhury et al., 2000; M. Wolf, M. Schreckenberg und A. Bachem, 1996), in der enzymatischen Biologie (Berry, 2003), in der Marktdynamik (Bouchaud and Georges, 1990).

²²Der von Prigogine und seiner Gruppe studierte Prototyp eines bestimmten Schemas chemischer Reaktionen, der Brüsselator, kann über eine Reihe von Bifurkationen fern vom Gleichgewicht ins Chaos getrieben werden. Diese Erkenntnisse wurden experimentell nachgewiesen, in der Biologie bei Zellvorgängen oder bei der Bildung von Termitenhügeln, aber auch bei sozialen Prozessen wie z. B. der Urbanisierung, wo die Stadt als eine dissipative Struktur angesehen wird. Prigogine propagierte "Dialog mit der Natur".

- Sprache (energy systems language, 1971) und mit Hilfe des Konzepts der Emergie (embodied or memorized energy, 1987) haben der amerikanische Ökologe Tom Odum und der australische Physiker und Mathematiker David M. Scienceman geprägt. Die Grundidee dabei war, in einem Netzwerk von gegenseitig nützlichen Beziehungen den (möglichen, aufgabenspezifischen) Nutzen eines “Produkts” anhand der zu seiner momentanen Gestaltung kumulativ benötigten nutzbaren Energie zu beschreiben. Der “Wert” von verschiedenartigen treibenden Prozessen kann dann z. B. einheitlich durch ein Sonnenenergie - Äquivalent ausgedrückt werden. Bei der Untersuchung des Energieflusses und seiner Transformationen (Energetik) wurde insbesondere vom Mathematiker und Chemiker Alfred J. Lotka (er ist auch für gleichnamige Populationsdynamik bekannt) und von H. T. Odum vorgeschlagen, das Prinzip der maximalen Leistung²³ als das vierte Prinzip der Thermodynamik anzuerkennen.

Die Physiker haben sich schon immer für Wachstumsprozesse (Kristallwachstum etc.) interessiert. Das Entdecken eines realistischen (Gleichgewicht -) Modells der diffusionsbegrenzten Anlagerung (DLA, 1981), in dem die Teilchen brownische Bewegung ausüben, bevor sie am Nukleationscluster hängen bleiben und dendritische Strukturen allmählich wachsen lassen, hat eine neue Welle der Forschungsaktivitäten über Wachstumsphänomene im (Nicht -) Gleichgewicht ausgelöst [Krug 97]. Die entsprechenden Modelle erfüllen die Voraussetzungen für einen robusten und relativ universellen Mechanismus, der zu einem komplexen Skalenverhalten führt.

Umfassende Studien über zelluläre Automaten (CA) betrieb in den 80er Jahren der englische Mathematiker Stephen Wolfram mit seinen Mitarbeitern [Wolf 86, Wolf 02]. Den Begriff hat er selbst eingeführt. In einem CA hängt der Zustand jedes Automaten in jedem Zeitschritt vom momentanen Zustand seiner “Nachbarn” ab. Die diskrete Natur der CA ist vorteilhaft, erlaubt jedoch keine dynamische Betrachtung mit Hilfe von Differentialgleichungen. Eine Kompromisslösung bilden gekoppelte Gitterabbildungen (Coupled Map Lattices, CML). Sie wurden Anfang der 80er Jahre eingeführt. Sie sind diskret in Raum und Zeit, werden wie CA durch lokale Korrelationen mit Nachbarn charakterisiert. Die lokalen Zustände werden aber durch kontinuierliche Felder dargestellt, so dass kritische Phasenübergänge zu beobachten sind [Boff 01]. Falls ein bestimmtes Verhalten als Ziel festgelegt ist, werden Simulationen zu diesen und anderen Modellen wie neuronale Netzwerke oder adaptive Systeme zunehmend mit Hilfe “kooperativer Agenten” (multi - agent system, MAS) durchgeführt, die lernfähig, anpassungsfähig, kooperativ etc. sein können.

Computersimulationen nehmen in der modernen Forschung über komplexe Systeme einen großen Stellenwert ein. Da der jeweilige Phasenraum meistens sehr groß ist, ist die vollständige Suche im Phasenraum (Berücksichtigung von allen möglichen Zuständen) selten praktikabel. Eine stochastische Suche im Phasenraum stellt z. B.

²³Für ein einem Selektionsdruck unterworfenen anpassungsfähiges System wird die auf die systemspezifische Produktion und Effizienz gerichtete Leistungsaufnahme und Energietransformation maximiert.

das Monte - Carlo - Verfahren (1953) dar, ein gängiges Mittel in der Physik. Beim Studium von zellulären Automaten wurde vom Psychologen und Elektrotechniker John H. Holland 1975 der Begriff genetischer Algorithmus eingeführt [Vose 99]. Dabei evolviert man ein MAS, indem zuerst eine Initialisierung der Individuen (Lösungskandidaten) und danach wiederholt Selektion der überlebenden Individuen anhand ihrer Fitness - Funktion und dann Erzeugung einer neuen Population mit Hilfe der Rekombination (zusätzliche neue Individuen werden aus den alten geschaffen) und Mutation (Individuen werden zufällig modifiziert) stattfindet. Bei der genetischen Programmierung werden dagegen nicht Individuen, sondern (auf eine Aufgabe spezialisierte) Programme selektiert. Die genetische Programmierung geht auf Stephen F. Smith (1980) und Michael L. Cramer (1985) zurück [Koza 92].

Das Konzept der Störungspropagation als eine Methode zur Untersuchung der dynamischen Antwort und anderer Eigenschaften von Netzwerken wurde im biologischen Kontext von Stuart Kauffman bereits 1969 vorgeschlagen [Kauf 69], ernsthaft eingesetzt wurde es jedoch erst seit der zweiten Hälfte der 80er Jahre und zwar in der Physik. Die Untersuchung der Propagation einer Störung muss mit Vorsicht angewandt werden, da die dynamische Entwicklung der Störung in den Untersuchungsobjekten, einer Systeminstanz aus dem Ensemble und ihrer Kopie mit der Störung, bekanntlich [Hirn 00] nicht nur von Modellparametern, sondern auch von den Details der jeweiligen Realisierung der Dynamik im System abhängen kann. Diese dynamische Entwicklung gibt nicht immer die durch (interessante) langreichweitige Korrelationen bestimmten Phasenränder wieder²⁴.

Das Studium der so genannten komplexen adaptiven Systeme (CAS) wurde seit Anfang der 90er Jahre am in Verbindung mit komplexen Systemen berühmten Santa Fe Institut vom Informatiker John H. Holland, dem berühmten Physiker Murray Gell - Mann, H. Morowitz, W. B. Arthur und anderen betrieben. Ein CAS ist eine makroskopische Sammlung von einfachen (gegebenenfalls klonierbaren) interagierenden Einheiten, die in ihrer dynamischen Umwelt anpassungsfähig sind und zusammen ein komplexes Verhalten zeigen. Der Unterschied zwischen CAS und MAS liegt in der größeren Komplexität der ersteren Systeme. In CAS beobachtet man räumliche und zeitliche Organisation, Selbstähnlichkeit, Selbstorganisation, Emergenz. Beispiele für CAS sind Modelle für Aktienmärkte, Kommunikation von sozialen Insekten wie in Ameisenkolonien, Unternehmen in einem Markt, Gruppen in sozialen Systemen, die Biosphäre und das Ökosystem, das Gehirn und das Immunsystem, die Zelle und die embryonale Entwicklung [Mang 01].

Netzwerke stellen komplexe Systeme auf eine natürliche Weise dar, wenn man die Rolle der Wechselbeziehungen im System erfassen will. (Dies ist die Essenz des Konnektionismus im Rahmen der allgemeinen Systemtheorie: Das komplexe Sys-

²⁴Ein einfaches Beispiel für die Diskrepanz: Wärmebad - Dynamik gegen Glauber - Dynamik für ein einziges Ising - Spin mit $\sigma = \pm 1$ bei $T \rightarrow \infty$. Bei der den Spin mit Wahrscheinlichkeit 1/2 in beiden Systemen umkehrenden Glauber - Dynamik wird die Störung für immer bestehen, während sie für die ausrichtende Wärmebad - Dynamik gleich im ersten Zeitschritt verschwindet. Ansonsten führen die beiden Dynamiken auf dasselbe Spinverhalten mit gleichwahrscheinlichen Werten $\sigma = \pm 1$.

temverhalten rührt von den Wechselbeziehungen her.) Bei einer Untersuchung der komplexen Netzwerke bedient man sich heutzutage der Nichtgleichgewichtstheorie, der Perkolationstheorie und vielem mehr. Da diese Theorien erst seit wenigen Jahrzehnten ein Gegenstand der aktiven Forschung sind, aber auch in Verbindung mit der rasanten Entwicklung der Computertechnologien und des Internets gedeiht die Forschung über komplexe Netzwerke erst seit den späten 90er Jahren. Besonders einflussreich waren dabei die Arbeiten von Stuart Kauffman über boolesche Netzwerke ab 1969 [Kauf 69], vom finnischen Ingenieur Teuvo Kohonen über neuronale Netzwerke in 1980ern [Koho 83], vom Soziologen Duncan Watts und dem Mechaniker Steven H. Strogatz über das kleine - Welt - Phänomen²⁵ in 1998 [Watt 98], vom rumänischen Physiker Albert - László Barabási und seiner Doktorandin Réka Albert über skalenfreie Topologien (bei denen die Verteilung der Knotengrade ein Potenzgesetz ist) in 1999 [Albe 99].

Viele neuere Modelle definieren eine bestimmte Dynamik auf den Netzwerken, also die Regeln, nach denen die verbundenen Knoten wechselwirken. Insbesondere Evolution und Wachstum von solchen Netzwerken, ihr Skalenverhalten, ihre Dynamik, gegebenenfalls ihre Selbstorganisation stehen im Mittelpunkt der Forschungsaktivitäten. Die Netzwerkmodelle stammen aus der Biologie, Soziologie, Physik, Ökologie, aus den Wirtschaftswissenschaften [Doro 02, Alda 03b].

Beim Studium von verschiedensten komplexen Systemen trifft man auf Parallelen, die im Rahmen der allgemeinen Systemtheorie erfasst werden können und auf potentielle Existenz von neuen Gesetzmäßigkeiten der Komplexität neben den bekannten wie die kritische Selbstorganisation oder verschiedene Erweiterungsversuche der thermodynamischen Grundsätze hinweisen.

²⁵Seit dem Milgram - Experiment in 1967, in dem Briefe von einer Testperson über nur wenige Zwischenstationen an ihr völlig unbekannte Adressaten weitergeleitet werden konnten, sind Netzwerke bekannt, die gleichzeitig einen großen Clustering - Koeffizienten und einen kleinen minimalen Abstand zwischen zwei Knoten haben.

B. Anzahl der Zyklen für zwei Schleifen mit Zwischenverbindung

B.1. Ergänzende Beispiele

Die Abbildungen B.1 und B.2 veranschaulichen die Ergebnisse des Abschn. 5.2.1.

B.2. Begründung der Formeln (5.8) und (5.9)

Hier werden die Formeln (5.8) und (5.9) genauer begründet. Sie stellen eine Näherung für die Anzahl der Zyklen im Fall $g > 1$ bei einer geraden ersten Schleife der Komponente mit zwei großen Schleifen und einer Zwischenverbindung aus dem Abschn. 5.2.1, g ist der größte gemeinsame Teiler von N_2 und der Länge des Zyklus der ersten Schleife P_1 .

Ich betrachte die Subsysteme aus Abschn. 5.2.1, in die die Knotenwerte von G_1 und G_2 während einer Periode des Zyklus zerfällt. Der große Unterschied in der Dynamik im Vergleich zum Fall der Primzahlen N_1 und N_2 ergibt sich dadurch, dass in einigen Subsystemen alle Werte von G_1 Null sein können ($G_1 = 0$), ich bezeichne sie als Nullsysteme. Nach Abschn. 5.2.1 versklavt die erste Schleife eines Subsystems die zweite, falls es kein Nullsystem ist, sonst befindet sich das Subsystem auf einem Zyklus der Länge $2N_2$ oder 2 (N_2 oder 1) für $f_\Sigma = f_{11}$ ($f_\Sigma = f_{14}$).

Ich bezeichne mit dem Index i die Anzahl der Nullsysteme, mit $\Theta(x)$ die Funktion, die für $x > 0$ Eins und sonst Null ist. Als erste wird die Funktion $f_\Sigma = f_{11}$ betrachtet. Da das kleinste gemeinsame Vielfache von N_1 und $2N_2$ für gerade N_1/g N_1N_2/g beträgt, ergeben sich zwei Fälle für die Anzahl der Zyklen der Länge $L = N_1N_2/g$ für gerade N_1/g und $L = 2N_1N_2/g$ für ungerade N_1/g :

$$C_{N_1, N_2}^{f_{11}, L} \simeq \frac{1}{L} \sum_{i=1}^{g-1} \binom{g}{i} ((2^{N_1/g} - 1)^{g-i} - 1) (2^{N_2 i/g} - \Theta(N_2/g - 2)2^i) = \quad (\text{B.1})$$

$$\frac{1}{L} \left[(2^{N_1/g} + 2^{N_2/g} - 1)^g - (2^{N_2/g} + 1)^g - \Theta(N_2/g - 2)(2^{N_1/g} + 1)^g \right. \\ \left. - (1 - \Theta(N_2/g - 2)) ((2^{N_1/g} - 1)^g - 1) + \Theta(N_2/g - 2)3^g \right] .$$

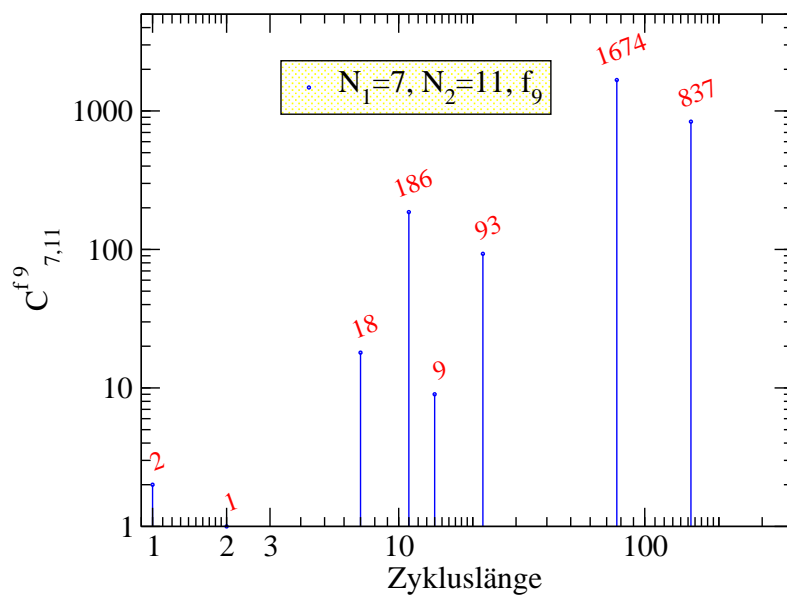


Abb. B.1.: Beispielhafte Darstellung der Anzahl der Zyklen für Schleifen mit Primzahlen - Größen und $f_\Sigma = f_9$.

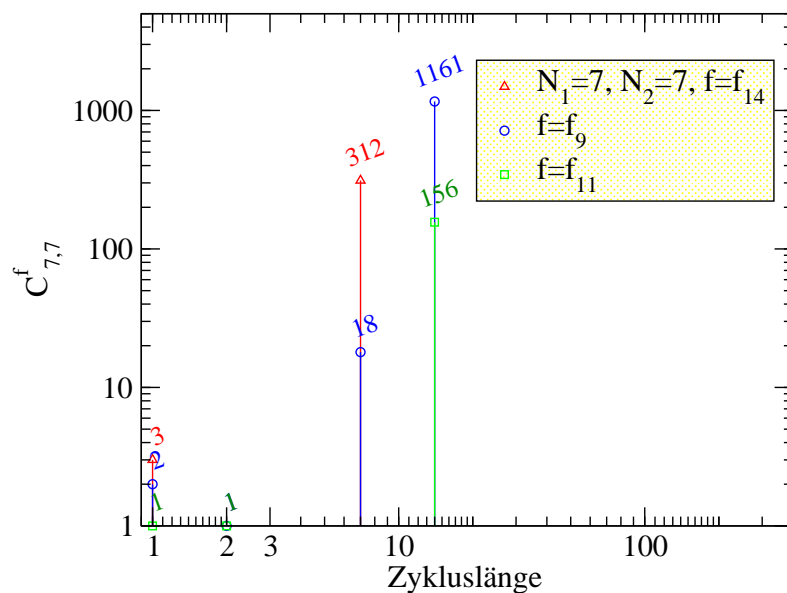


Abb. B.2.: Beispielhafte Darstellung der Anzahl der Zyklen für den “resonanten” Fall $N_1 = N_2$.

Der Term $\Theta(N_2/g - 2)2^i$ beschreibt die Anzahl der Zustände der Menge der zweiten Schleifen der Nullsysteme auf gemeinsamen Zyklen der Länge 2. Die Subtraktion der zweiten 1 im zweiten Faktor unter der Summe berücksichtigt den Fall, dass alle Knotenwerte der ersten Schleife aus den Nichtnullsystemen 1 sind. Diese Subtraktion muss im Spezialfall $N_1 = 2g$ im Faktor fallen gelassen werden. Falls N_1/g oder N_2/g keine Primzahlen sind, ist zu beachten, dass kleine Zyklen mit Längen, die Teiler von L sind, hier in (B.1) und auch weiter unten in diesem Abschnitt vernachlässigt werden.

Die Anzahl der Zyklen der Länge N_1 beträgt für gerade N_1/g

$$\begin{aligned} C_{\text{even}, N_1, N_2}^{f_{11}, N_1} &\simeq \frac{1}{N_1} \sum_{i=0}^{(g-1)\Theta(\frac{N_2}{g}-2)} \binom{g}{i} ((2^{N_1/g} - 1)^{g-i} - 1)(1 + \Theta(\frac{N_2}{g} - 2))^i \quad (\text{B.2}) \\ &= \frac{1}{N_1} [(2^{N_1/g} + 2\Theta(N_2/g - 2) - 1)^{(g-1)\Theta(N_2/g-2)+1} - (1 + 2\Theta(N_2/g - 2))^g] . \end{aligned}$$

Im Spezialfall $N_1 = 2g$ muss die Summe bis $i = g$ gehen und die Subtraktion der zweiten 1 im zweiten Faktor unter der Summe muss fallen gelassen werden.

Bei geraden (ungeraden) N_1/g und $N_2/g > 2$ treten Zyklen der Länge $L = N_1$ ($L = 2N_1$) auf, die durch Kombination von Subsystemen auf einem Zyklus der Länge N_1/g und von Nullsystemen auf einem Zyklus der Länge 2 entstehen. Ihre Anzahl ist

$$C_{N_1, N_2}^{f_{11}, L} \simeq \frac{1}{2N_1} \sum_{i=1}^{g-1} \binom{g}{i} ((2^{N_1/g} - 1)^{g-i} - 1)2^i . \quad (\text{B.3})$$

Durch Kombination von Subsystemen auf einem Zyklus der Länge N_1/g und von Fixpunkt - Subsystemen gibt es eine zusätzliche Anzahl der Zyklen der Länge N_1 :

$$C_{N_1, N_2}^{f_{11}, N_1} \simeq \frac{1}{N_1} ((2^{N_1/g} - 1)^g - 1) . \quad (\text{B.4})$$

Die Anzahl der Zyklen der Länge $2N_2$ beträgt

$$\begin{aligned} C_{N_1, N_2}^{f_{11}, 2N_2} &\simeq \frac{1}{2N_2} \sum_{i=1}^g \binom{g}{i} (2^{N_2 i/g} - \Theta(N_2/g - 2)2^i) = \quad (\text{B.5}) \\ &\frac{1}{2N_2} [(2^{N_2/g} + 1)^g - 1 - \Theta(N_2/g - 2)(3^g - 1)] \end{aligned}$$

Als Nächstes wird $f_\Sigma = f_{14}$ betrachtet. Die Anzahl der Zyklen der Länge $L = N_1 N_2/g$ erhält man aus (B.1), indem man dort alle Θ - Funktionen durch 1 ersetzt. Dazu kommen die Zyklen der Länge N_1 :

$$C_{N_1, N_2}^{f_{14}, N_1} \simeq \frac{1}{N_1} \sum_{i=0}^{g-1} \binom{g}{i} ((2^{N_1/g} - 1)^{g-i} - 1)2^i = \frac{1}{N_1} [(2^{N_1/g} + 1)^g - 3^g] . \quad (\text{B.6})$$

B. Anzahl der Zyklen für zwei Schleifen mit Zwischenverbindung

Die Zyklen der Länge N_2 ergeben sich schließlich, wenn es ein oder mehr Subsysteme mit allen Knotenwerten 0 gibt und alle Knotenwerte der restlichen Subsysteme 1 sind. Die Anzahl der Zyklen der Länge N_2 beträgt

$$C_{N_1, N_2}^{f_{14}, N_2} \simeq \frac{1}{N_2} \sum_{i=1}^g \binom{g}{i} (2^{N_2 i/g} - 2^i) = \frac{1}{N_2} [(2^{N_2/g} + 1)^g - 3^g] . \quad (\text{B.7})$$

Durch Zusammenfassung der Ergebnisse (B.1) bis (B.7) ergibt sich schnell die folgende Näherung für die Anzahl der Zyklen für kanalisierende Funktionen f_Σ (gerade erste Schleife, große Schleifen, $g > 1$, $\alpha \equiv (N_1/g \bmod 2 + 1)$):

$$\begin{aligned} C_{N_1, N_2}^{f_{14}} &\simeq \frac{g(2^{N_1/g} + 2^{N_2/g} - 1)^g}{N_1 N_2} + \frac{(2^{N_1/g} + 1)^g - 3^g}{N_1} + \frac{(2^{N_2/g} + 1)^g - 3^g}{N_2} , \\ C_{N_1, N_2}^{f_{11}} &\simeq \frac{g(2^{N_1/g} + 2^{N_2/g} - 1)^g}{\alpha N_1 N_2} + \frac{(2^{N_1/g} + 1)^g + (\alpha - 1)(2^{N_1/g} - 1)^g - 3^g}{\alpha N_1} \\ &\quad + \frac{(2^{N_2/g} + 1)^g - 3^g}{2N_2} . \end{aligned} \quad (\text{B.8})$$

Die Formeln (5.8) und (5.9) folgen bei großen N_1 und N_2 .

C. Ergänzende Ergebnisse für Schleifen mit einer Querverbindung

Diese allgemein gültigen Ergebnisse für die Anzahl und Länge der Zyklen der Schleifen mit einer Querverbindung ergänzen diejenigen aus Abschn. 5.2.2.

Als erstes wird eine untere Schranke für die Anzahl der Zyklen einer Schleife mit einer Querverbindung mit festen N und L präsentiert. Für gerade N und L zerfällt die Menge der Knoten in zwei unabhängige Subsysteme. Da zwei gegebene Zyklen der zwei gleichen Subsysteme verschiedene Phasendifferenzen haben können, wenn ihre Längen gemeinsame Teiler besitzen, gilt die Ungleichung

$$C_{2N}^{2L} \geq C_N^L * C_N^L / 2.$$

Mehrfache Anwendung führt auf

$$C_{2^\nu N_0}^{2^\nu L_0} \geq \left(C_{N_0}^{L_0} / \sqrt{2} \right)^{2^\nu} \equiv C_0^{2^\nu N_0} = C_0^N.$$

Da jeder Wert von L bei festem N mit der Wahrscheinlichkeit $1/N$ auftritt, erfüllt die mittlere Zyklenanzahl der Komponente der Größe $N = 2^\nu N_0$ die Ungleichung

$$C_N \geq \frac{1}{N} (C_0)^N \equiv 2^{AN} / N. \quad (\text{C.1})$$

Die Zyklenanzahl steigt exponentiell mit N an.

Die nächste Bemerkung betrifft die Tatsache, dass in numerischen Simulationen durchschnittlich immer ein Fixpunkt pro Komponente gefunden wird.

Im Ensemble der Komponenten der Größe N mit allen möglichen Kombinationen der Knotenfunktionen ist der Nachfolger eines Komponentenzustands im Zustandsraum mit der gleichen Wahrscheinlichkeit ein beliebiger Zustand, inklusive des ursprünglichen Zustands [Samu 03]. Die Wahrscheinlichkeit für einen Fixpunkt beträgt daher $1/2^N$. Die Summation über alle Zustände liefert im Mittel einen Fixpunkt pro Komponente.

Schließlich wird erklärt, warum im Fall der kanalisierenden Funktionen mit Wahrscheinlichkeit $1/4$ ein Zyklus der Länge 2 auftritt.

Für reversible f_Σ gibt es keine Zyklen der Länge 2. Je nach L würden die Eingänge von Σ auf einem Zyklus der Länge 2 entweder alternierend die Werte $(0, 1)$ und $(1, 0)$ oder alternierend die Werte $(0, 0)$ and $(1, 1)$ annehmen. In beiden Fällen wäre

der Ausgang der reversiblen Funktion konstant im Widerspruch zur Annahme des Zyklus der Länge 2.

Für kanalisierende Funktionen f_Σ werden zwei verschiedene Beweise für die Wahrscheinlichkeit $1/4$ gegeben.

Für einen Komponentenzustand $g^{(i)}$ ist die Wahrscheinlichkeit, dass auf $g^{(i)}$ ein anderer Zustand $g^{(j)}$ folgt, analog zur Betrachtung der Fixpunkte $1/2^N$ [Samu 03]. Ich bezeichne die Menge der Komponenten, in denen dieser Übergang vorkommt mit \mathcal{N}_{ij} . Wie groß ist die Wahrscheinlichkeit, dass der Zustand $g^{(j)}$ im nächsten Zeitschritt wieder zu $g^{(i)}$ wird? Für die Komponenten aus \mathcal{N}_{ij} muss f_Σ einen bestimmten Wert für einen der 4 Eingangszustände annehmen um den Übergang $i \mapsto j$ für feste i und j leisten zu können. Daher kommen 4 der 8 kanalisierenden Funktionen für f_Σ nicht in Frage und die Wahrscheinlichkeit, dass eine Komponente aus \mathcal{N}_{ij} den Übergang von $g^{(j)}$ nach $g^{(i)}$ leistet beträgt $4/(8 * 2^N)$. Zusammenfassend ist die Wahrscheinlichkeit p_2 eines Zyklus der Länge 2:

$$p_2 = \frac{1}{2} \sum_{i,j} \frac{1}{2^N} \frac{1}{2 * 2^N} = 1/4 . \quad (\text{C.2})$$

Das kann man auch durch eine explizite Konstruktion der Zyklen der Länge 2 beweisen. Diese Zyklen sind Folgen der alternierenden Nullen und Einsen, die für ungerade N am Platz Σ ausnahmsweise zwei Nullen oder zwei Einsen beisammen stehen haben. (Ohne Beschränkung der Allgemeinheit sind alle Knotenfunktionen außer f_Σ Kopierfunktionen.) Σ reagiert entweder auf alternierende Eingangszustände 01 und 10 für ungerade L oder auf alternierende Eingangszustände 00 und 11 für gerade L , der Ausgang muss dabei abwechselnd 0 und 1 sein. In jedem Fall, das heißt, für alle festen N und L , erfüllen zwei der acht kanalisierenden Funktionen diese Bedingung. Der Anteil der Komponenten, die einen Zyklus der Länge 2 haben, beträgt für alle L und alle Knotenfunktionen $2/8 = 1/4$. Ein Beispiel: Für ungerade L und ungerade N muss der Ausgang für den Eingangszustand 01 (das rechte Bit gehört dem ersten Eingang G_1) 1 sein; die Vorgabe für 10 ist 0. Diese Vorgaben erfüllen zwei kanalisierende Funktionen.

Literaturverzeichnis

- [Ahl 96] V. Ahl and T. F. H. Allen. *Hierarchy theory, a vision, vocabulary and epistemology*. Columbia University Press, 1996.
- [Akut 00] T. Akutsu, S. Miyano, and S. Kuhara. “Inferring qualitative relations in genetic networks and metabolic pathways”. *Bioinformatics*, Vol. 16, pp. 727–734, 2000.
- [Akut 99] T. Akutsu, S. Miyano, and S. Kuhara. “Identification of genetic networks from a small number of gene expression patterns under the Boolean network model”. In: *Proc. Pacific Symposium on Biocomputing (PSB99)*, pp. 17–28, 1999.
- [Albe 03] R. Albert and H. G. Othmer. “The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*”. *J. Theor. Biol.*, Vol. 223, pp. 1–18, 2003.
- [Albe 05] R. Albert. “Scale-free networks in cell biology”. *Journal of cell science*, Vol. 118, pp. 4947–4957, 2005.
- [Albe 99] R. Albert, H. Jeong, and A.-L. Barabási. “The diameter of the world-wide web”. *Nature*, Vol. 401, p. 130, 1999.
- [Alda 03a] M. Aldana. “Dynamics of Boolean Networks with Scale Free Topology”. *Physica D*, Vol. 185, No. 1, pp. 45–66, 2003.
- [Alda 03b] M. Aldana-Gonzalez, S. Coppersmith, and L. P. Kadanoff. “Boolean Dynamics with Random Couplings”. In: E. Kaplan, J. E. Marsden, and K. R. Sreenivasan, Eds., *Perspectives and Problems in Nonlinear Science, A celebratory volume in honor of Lawrence Sirovich*, pp. 23–89, Springer Applied Mathematical Sciences Series, Springer Verlag, New York, 2003.
- [Alon 99] U. Alon, M. G. Surette, N. Barkai, and S. Leibler. “Robustness in bacterial chemotaxis”. *Nature*, Vol. 397, pp. 168–171, 1999.
- [Andr 05] M. Andrecut. “Mean field dynamics of random Boolean networks”. *J. Stat. Mech. (JSTAT)*, Vol. P02003, 2005.
- [Andr 06] M. Andrecut and S. A. Kauffman. “Mean-field model of genetic regulatory networks”. *New J. Phys.*, Vol. 8, p. 148, 2006.

- [Anta 97] T. Antal, Z. Rácz, and L. Sasvári. “Nonequilibrium Steady State in a Quantum System: One-Dimensional Transverse Ising Model with Energy Current”. *Phys. Rev. Lett.*, Vol. 78, p. 167, 1997.
- [Ashb 47] W. R. Ashby. “Principles of the Self-Organizing Dynamic System”. *Journal of General Psychology*, Vol. 37, pp. 125–128, 1947.
- [Babu 04] M. Babu, N. Luscombe, L. Aravind, M. Gerstein, and S. Teichmann. “Structure and evolution of transcriptional regulatory networks”. *Curr. Opin. Struct. Biol.*, Vol. 14, No. 3, pp. 283–291, 2004.
- [Bar 03] Y. Bar-Yam. *Making Things Work: Solving Complex Problems in a Complex World*. NECSI, Knowledge Press, U.S., 2003.
- [Bast 96] U. Bastolla and G. Parisi. “Closing Probabilities in the Kauffman Model: an Annealed Computation”. *Physica D*, Vol. 98, pp. 1–25, 1996.
- [Bast 98a] U. Bastolla and G. Parisi. “The Modular Structure of Kauffman Networks”. *Physica D*, Vol. 115, No. 3&4, p. 219, 1998.
- [Bast 98b] U. Bastolla and G. Parisi. “Relevant elements, magnetization and dynamical properties in Kauffman networks. A numerical study”. *Physica D*, Vol. 115, No. 3&4, p. 203, 1998.
- [Bert 45] K. L. von Bertalanffy. “Zu einer allgemeinen Systemlehre”. *Blätter für Deutsche Philosophie*, Vol. 195, No. 3/4, 1945.
- [Bert 54] K. L. von Bertalanffy. *General System theory: Foundations, Development, Applications*. George Braziller New York; Revised edition, 1976 (1954).
- [Bigg 98] J. Biggart, G. Glovelli, and A. Yassour. *Bogdanov and His Work: A Guide to the Published and Unpublished Works of Alexander A. Bogdanov (Malinovsky) 1873-1928*. Aldershot, England and Brookfield, Vermont: Ashgate, 1998.
- [Bilk 02] S. Bilke and F. Sjunnesson. “Stability of the Kauffman model”. *Phys. Rev. E*, Vol. 65, p. 016129, 2002.
- [Boff 01] G. Boffetta, M. Genchini, M. Falcioni, and A. Vulpiani. “Predictability: a way to characterize complexity”. *Physics Reports*, Vol. 356, pp. 367–474, 2001.
- [Bohm 87] D. Bohm and F. D. Peat. *Science, Order & Creativity*. Bantam Books, New York, 1987.
- [Born 00] S. Bornholdt and T. Rohlf. “Topological Evolution of Dynamical Networks: Global Criticality from Local Dynamics”. *Phys. Rev. Lett.*, Vol. 84, No. 26, pp. 6114–6117, 2000.
- [Bowe 01] J. M. Bower and H. Bolouri. *Computational Modeling of Genetic and Biochemical Networks*. MIT Press, Cambridge, MA, 2001.

- [Bran 05] O. Brandman, J. E. Ferrell Jr., R. Li, and T. Meyer. “Interlinked Fast and Slow Positive Feedback Loops Drive Reliable Cell Decisions”. *Science*, Vol. 310, pp. 496–498, 2005.
- [Brau 06] S. Braunewell and S. Bornholdt. “Superstability of the yeast cell cycle dynamics: Ensuring causality in the presence of biochemical stochasticity”. 2006. q-bio.MN/0605009.
- [Braz 03] A. Brazma and T. Schlitt. “Reverse engineering of gene regulatory networks: a finite state linear model”. *Genome Biology*, Vol. 4, p. 5, 2003.
- [Broa 57] S. R. Broadbent and J. M. Hammersley. “Percolation processes I. Crystals and mazes”. In: *Proc. Camb. Phil. Soc.*, pp. 629–641, 1957.
- [Capr 00] F. Capra. *The Tao of Physics*. Shambala, 4th edition, 2000.
- [Chav 05] M. Chaves, R. Albert, and E. D. Sontag. “Robustness and fragility of Boolean models for genetic regulatory networks”. *J. Theor. Biol.*, Vol. 235, No. 3, pp. 431–449, 2005.
- [Chav 06] M. Chaves, E. D. Sontag, and R. Albert. “Methods of robustness analysis for Boolean models of gene control networks”. In: *IEE Proc. Systems Biology*, pp. 154–167, 2006.
- [Chen 02] L. Chen and K. Aihara. “Stability of Genetic Regulatory Networks With Time Delay”. *IEEE Transactions on Circuits and Systems I*, Vol. 49, No. 5, p. 602, 2002.
- [Chen 99] T. Chen, H. L. He, and G. M. Church. “Modelling gene expression with differential equations”. In: *Proc. Pacific Symposium on Biocomputing (PSB99)*, pp. 29–40, 1999.
- [Cohe 02] R. Cohen, D. ben Avraham, and S. Havlin. “Percolation critical exponents in scale-free networks”. *Phys. Rev. E*, Vol. 66, p. 036113, 2002.
- [Corn 83] P. A. Corning. *The Synergism Hypothesis: A Theory of Progressive Evolution*. New York: McGraw-Hill, 1983.
- [Cros 93] M. Cross and P. Hohenberg. “Pattern formation outside of equilibrium”. *Rev. Mod. Phys.*, Vol. 65, pp. 851–1112, 1993.
- [Dalt 01] F. Dalton and D. Corcoran. “Self-organized criticality in a sheared granular stick-slip system”. *Phys. Rev. E*, Vol. 63, p. 061312, 2001.
- [Dass 00] G. von Dassow, E. Meir, E. M. Munro, and G. M. Odell. “The segment polarity network is a robust developmental module”. *Nature*, Vol. 406, pp. 188–192, 2000.
- [Davi 02] E. H. Davidson, J. P. Rast, P. Oliveri, A. Ransick, C. Caletani, C. H. Yuh, T. Minokawa, G. Amore, V. Hinman, C. Arenas-Mena, and et al. “A Genomic Regulatory Network for Development”. *Science*, Vol. 295, pp. 1669–1678, 2002.

- [Dee 84] D. Dee and M. Ghil. “Boolean Difference Equations I: Formulation and Dynamic Behavior”. *SIAM Journal on Applied Mathematics*, Vol. 44, No. 1, pp. 111–126, 1984.
- [Deok 06] L. Deok-Sun and H. Rieger. “Broad edge of chaos in strongly heterogeneous Boolean networks”. 2006. cond-mat/0605730.
- [Derr 86a] B. Derrida and Y. Pomeau. “Random networks of automata: a simple annealed approximation”. *Europhys. Lett.*, Vol. 1, No. 2, pp. 45–49, 1986.
- [Derr 86b] B. Derrida and D. Stauffer. “Phase transitions in two dimensional Kauffman cellular automata”. *Europhys. Lett.*, Vol. 2, No. 10, pp. 739–745, 1986.
- [Derr 87] B. Derrida. “Dynamical Phase Transition in Non-Symmetric Spin Glasses”. *J. Phys. A*, Vol. 20, pp. L721–L725, 1987.
- [Desc 85] R. Descartes. *The Philosophical Writings of Descartes*. Vol. 1, Cambridge University Press, 1985.
- [Dhae 00] P. D’haeseleer, S. Liang, and R. Somogyi. “Genetic network inference: from co-expression clustering to reverse engineering”. *Bioinformatics*, Vol. 16, No. 8, 2000.
- [Dhae 99] P. D’haeseleer, X. Wen, S. Fuhman, and R. Somogyi. “Linear modeling of mRNA expression levels during CNS development and injury”. In: *Proc. Pacific Symp. on Biocomputing*, pp. 41–52, 1999.
- [Di P 01] E. A. Di Paolo. “Rhythmic and non-rhythmic attractors in asynchronous random Boolean networks”. *BioSystems*, Vol. 59, pp. 185–195, 2001.
- [Dick 02] R. Dickman. “Nonequilibrium Phase Transitions in Epidemics and Sandpiles”. *Physica A*, Vol. 306, p. 90, 2002.
- [Doro 02] S. Dorogovtsev and J. Mendes. “Evolution of networks”. *Adv. Phys.*, Vol. 51, p. 1079, 2002.
- [Dros 05a] B. Drossel. “Number of attractors in random Boolean networks”. *Phys. Rev. E*, Vol. 72, p. 016110, 2005.
- [Dros 05b] B. Drossel, T. Mihaljev, and F. Greil. “Number and length of attractors in a critical Kauffman model with connectivity one”. *Phys. Rev. Lett.*, Vol. 94, p. 088701, 2005.
- [Dudl 96] P. Dudley. “Back to basics? Tektology and General System Theory (GST)”. *Systemic Practice and Action Research*, Vol. 9, No. 3, pp. 273–284, 1996.
- [Edwa 01] R. Edwards, H. T. Siegelmann, K. Aziza, and L. Glass. “Symbolic dynamics and computation in model gene networks”. *Journal of Chaos*, Vol. 11, No. 1, pp. 160–169, 2001.
- [Eins 54] A. Einstein. *Ideas and Opinions*. New York, Crown Publishers, 1954.

- [Flyv 88a] H. Flyvbjerg. “An order parameter for networks of automata”. *J. Phys. A*, Vol. 21, p. L955, 1988.
- [Flyv 88b] H. Flyvbjerg and N. J. Kjær. “Exact solution of Kauffman’s model with connectivity one”. *J. Phys. A*, Vol. 21, p. 1695, 1988.
- [Galp 04] M. Y. Galperin. “Bacterial signal transduction network in a genomic perspective”. *Environ. Microbiol.*, Vol. 6, No. 6, pp. 552–567, 2004.
- [Gard 00] T. S. Gardner, C. R. Cantor, and J. J. Collins. “Construction of a genetic toggle switch in *Escherichia Coli*”. *Nature*, Vol. 403, No. 20, pp. 339–342, 2000.
- [Gard 90] C. W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*. Springer, Berlin, 1990.
- [Gat 03] I. Gat-Viks and R. Shamir. “Chain functions and scoring functions in genetic networks”. *Bioinformatics*, Vol. 19, No. Suppl. 1, pp. i108–i117, 2003.
- [Gers 04] C. Gershenson. “Updating schemes in Random Boolean Networks: do they really matter?”. In: J. Pollack, M. Bedau, P. Husbands, T. Ikegami, and R. A. Watson, Eds., *Artificial Life IX, Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems*, pp. 238–243, MIT Press, 2004.
- [Glas 73] L. Glass and S. A. Kauffman. “The logical analysis of continuous, non-linear biochemical control networks”. *J. Theor. Biol.*, Vol. 39, pp. 103–129, 1973.
- [Glas 75] L. Glass. “Classification of biological networks by their qualitative dynamics”. *J. Theor. Biol.*, Vol. 54, pp. 85–107, 1975.
- [Glas 78] L. Glass and J. S. Pasternak. “Prediction of limit cycles in mathematical models of biological oscillations”. *Bull. Math. Biol.*, Vol. 40, pp. 27–44, 1978.
- [Gold 99a] N. Goldenfeld, L. P. Kadanoff, G. M. Whitesides, R. F. Ismagilov, G. Weng, U. S. Bhalla, R. Iyengar, D. Rind, and W. B. Arthur. “Complex Systems: Viewpoints”. *Science*, Vol. 284, No. 5411, pp. 87–109, 1999.
- [Gold 99b] J. Goldstein. “Emergence as a construct: History and issues”. *Emergence*, Vol. 1, 1999.
- [Gonz 06] A. Gonzalez, A. Naldi, L. Sánchez, D. Thieffry, and C. Chaouiya. “GINsim: a software suite for the qualitative modelling, simulation and analysis of regulatory networks”. *Biosystems*, Vol. 84, No. 2, pp. 91–100, 2006.
- [Goss 98] P. J. E. Goss and J. Peccoud. “Quantitative modeling of stochastic systems in molecular biology by using stochastic Petri nets”. *Proc. Nat. Acad. Sci.*, Vol. 95, pp. 6750–6755, 1998.
- [Gras 59] P. Grassé. “La reconstruction du nid et les interactions inter-individuelles chez les bellicositermes natalenis et cubitermes sp. La théorie de la stigmergie: essai d’interprétation des termites constructeurs”. *Insectes Sociaux*, Vol. 6, pp. 41–83, 1959.

- [Gref 06] J. Grefenstette, S. Kima, and S. Kauffman. “An analysis of the class of gene regulatory functions implied by a biochemical model”. *BioSystems*, Vol. 84, pp. 81–90, 2006.
- [Grei 05] F. Greil and B. Drossel. “The dynamics of critical Kauffman networks under asynchronous stochastic update”. *Phys. Rev. Lett.*, Vol. 95, p. 048701, 2005.
- [Hake 83] H. Haken. *Synergetics: An Introduction. Nonequilibrium Phase Transition and Self-Organization in Physics, Chemistry, and Biology, Third Revised and Enlarged Edition*. Springer-Verlag, 1983.
- [Harr 02] S. Harris, B. Sawhill, A. Wuensche, and S. Kauffman. “A model of transcriptional regulatory networks based on biases in the observed regulation rules”. *Complexity*, Vol. 7, No. 4, pp. 23–40, 2002.
- [Hart 40] N. Hartmann. *Der Aufbau der realen Welt : Grundriß d. allg. Kategorienlehre*. de Gruyter Berlin, 1940.
- [Harv 97] I. Harvey and T. Bossomaier. “Time out of joint: attractors in asynchronous random Boolean networks”. In: P. Husbands and I. Harvey, Eds., *Processings of the Fourth European Conference on Artificial Life*, pp. 67–75, MIT Press, Cambridge, MA, 1997.
- [Hinr 00] H. Hinrichsen. “Nonequilibrium Critical Phenomena and Phase Transitions into Absorbing States”. *Adv. Phys.*, Vol. 49, pp. 815–958, 2000.
- [Hira 02] H. Hirata and et al. “Oscillatory expression of the bHLH Factor Hes1 regulated by a negative feedback loop”. *Science*, Vol. 298, pp. 840–843, 2002.
- [Hohe 77] P. C. Hohenberg and B. I. Halperin. “Theory of dynamic critical phenomena”. *Rev. Mod. Phys.*, Vol. 49, pp. 435–479, 1977.
- [Huan 99] S. Huang. “Gene expression profiling, genetic networks, and cellular states: an integrating concept for tumorigenesis and drug discovery”. *J. Mol. Med.*, Vol. 77, pp. 469–480, 1999.
- [Jaco 61] F. Jacob and J. Monod. “Genetic regulatory mechanisms in the synthesis of proteins”. *J. Molecular Biol.*, Vol. 3, pp. 318–356, 1961.
- [Jens 89] H. J. Jensen, K. Christensen, and H. C. Fogedby. “ $1/f$ noise, distribution of lifetimes, and a pile of sand”. *Phys. Rev. B*, Vol. 40, p. 7425, 1989.
- [Jeon 00] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A. L. Barabasi. “The large scale organization of metabolic networks”. *Nature*, Vol. 407, pp. 651–654, 2000.
- [Jong 02] H. de Jong. “Modeling and simulation of genetic regulatory systems: A literature review”. *J. Comp. Biol.*, Vol. 9, No. 1, pp. 69–105, 2002.
- [Jong 04] H. de Jong, J.-L. Gouzé, C. Hernandez, M. Page, T. Sari, and J. Geiselman. “Qualitative simulation of genetic regulatory networks using piecewise-linear models”. *Bull. Math. Biol.*, Vol. 66, pp. 301–340, 2004.

- [Jong 06] H. de Jong, C. Chaouiya, and T. D. “Dynamical modeling of biological regulatory networks”. *BioSystems*, Vol. 84, pp. 77–80, 2006.
- [Just 04] W. Just, I. Shmulevich, and J. Konvalina. “The number and probability of canalizing functions”. *Physica D*, Vol. 197, pp. 211–221, 2004.
- [Kaln 06] G. Kalna and D. J. Higham. “Clustering Coefficients for Weighted Networks”. *University of Strathclyde Mathematics Research report*, Vol. 3, 2006. Submitted to AISB’06 - Adaptation in Artificial and Biological Systems Proceedings.
- [Kamp 92] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier Science B.V., Amsterdam, 1992.
- [Kauf 03] S. Kauffman, C. Peterson, B. Samuelsson, and C. Troein. “Random Boolean network models and the yeast transcriptional network”. *Proc. Nat. Acad. Sci.*, Vol. 100, p. 14796, 2003.
- [Kauf 05a] V. Kaufman and B. Drossel. “On the properties of cycles of simple Boolean networks”. *Eur. Phys. J. B*, Vol. 43, pp. 115–124, 2005.
- [Kauf 05b] V. Kaufman, T. Mihaljev, and B. Drossel. “Scaling in critical random Boolean networks”. *Phys. Rev. E*, Vol. 72, p. 046124, 2005.
- [Kauf 06] V. Kaufman and B. Drossel. “Relevant components in critical random Boolean networks”. *New J. Phys.*, Vol. 8, p. 228, 2006.
- [Kauf 69] S. A. Kauffman. “Metabolic Stability and Epigenesis in Randomly Constructed Genetic Nets”. *J. Theor. Biol.*, Vol. 22, pp. 437–467, 1969.
- [Kauf 74] S. Kauffman. “The large scale structure and dynamics of gene control circuits: an ensemble approach”. *J. Theor. Biol.*, Vol. 44, pp. 167–190, 1974.
- [Kauf 85] M. Kauffman, J. Urbain, and R. Thomas. “Towards a logical analysis of immune response”. *J. Theor. Biol.*, Vol. 114, pp. 527–561, 1985.
- [Kauf 93] S. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, 1993.
- [Kauf 99] M. Kaufman, F. Andris, and O. Leo. “A logical analysis of T cell activation and anergy”. *Proc. Nat. Acad. Sci.*, Vol. 96, pp. 3894–3899, 1999.
- [Kita 04] H. Kitano. “Biological robustness”. *Nature Reviews Genetics*, Vol. 5, pp. 826–837, 2004.
- [Klem 05a] K. Klemm and S. Bornholdt. “Stable and unstable attractors in Boolean networks”. *Phys. Rev. E*, Vol. 72, p. 055101, 2005.
- [Klem 05b] K. Klemm and S. Bornholdt. “Topology of biological networks and reliability of information processing”. *Proc. Nat. Acad. Sci.*, Vol. 102, No. 51, pp. 18414–18419, 2005.

- [Koho 83] T. Kohonen. *Self-Organisation and Associative Memory*. Springer, Berlin, 1983.
- [Korz 95] A. Korzybski. *Science and Sanity - An Introduction to Non-Aristotelian Systems and General Semantics, 5th edition*. Inst. of General Semantics, 1995.
- [Koza 92] J. R. Koza. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. The MIT Press, 1992.
- [Krug 97] J. Krug. “Origins of scale invariance in growth processes”. *Advances in Physics*, Vol. 46, No. 2, pp. 139–282, 1997.
- [Küpf 49] K. Küpfmüller. *Die Systemtheorie der elektrischen Nachrichtenübertragung*. Stuttgart: Hirzel, 3. Aufl., 1968(1949).
- [Lago 05] M. C. Lagomarsino, P. Jona, and B. Bassetti. “The large-scale logico-chemical structure of a transcriptional regulation network”. 2005. q-bio.MN/0502017.
- [Laoz 06a] Laozi. “Tao Te King”. Übersetzt von Richard Wilhelm <http://gutenberg.spiegel.de/wilhelm/taotekin/taotekin.htm>, 2006.
- [Laoz 06b] Laozi. “Tao Te King”. Auszug von Tom Mandel, interpretiert von V. Kaufman, http://www.worldtrans.org/ISSS_Primer/seminare.html, 2006.
- [Laub 04] R. Laubenbacher and B. Stigler. “A Computational Algebra Approach to the Reverse Engineering of Gene Regulatory Networks”. *J. Theor. Biol.*, Vol. 229, pp. 523–537, 2004.
- [Lee 02] T. I. Lee, N. J. Rinaldi, F. Robert, D. T. Odom, and Z. Bar-Joseph. “Transcriptional regulatory networks in *saccharomyces cerevisiae*”. *Science*, Vol. 298, pp. 799–804, 2002.
- [Leon 51] W. W. Leontief. *Input-Output Economics*. Scientific American, 1951.
- [Levi 05] M. Levine and E. H. Davidson. “Gene regulatory networks for development”. *Proc. Nat. Acad. Sci.*, Vol. 102, pp. 4936–4942, 2005.
- [Lewe 75] G. Lewes. *Problems of Life and Mind. Volume 2*. Kegan Paul, Trench, Turbner, London, 1875.
- [Li 04] F. Li, T. Long, Y. Lu, Q. Ouyang, and C. Tang. “The yeast cell-cycle network is robustly designed”. *Proc. Nat. Acad. Sci.*, Vol. 101, No. 14, pp. 4781–4786, 2004.
- [Lian 98] S. Liang, S. Fuhrman, and R. Somogyi. “REVEAL, A General Reverse Engineering Algorithm for Inference of Genetic Network Architectures”. In: *Proc. Pacific Symposium on Biocomputing*, pp. 18–29, 1998.
- [Lusc 04] N. M. Luscombe, M. M. Babu, H. Y. Yu, M. Snyder, S. S. Teichmann, and M. Gerstein. “Genomic analysis of regulatory network dynamics reveals large topological changes”. *Nature*, Vol. 431, No. 308, pp. 308–312, 2004.

- [Mand 82] B. B. Mandelbrot. *The Fractal Geometry of Nature*. New York: W. H. Freeman and Co., 1982.
- [Mang 01] M. Mangel. “Complex Adaptive Systems, Aging and Longevity”. *J. Theor. Biol.*, Vol. 213, pp. 559–571, 2001.
- [Mars 84] G. Marsaglia. “A Current View of Random Number Generators”. In: *Computer Science and Statistics: Proceedings of the Symposium on the Interface, 16th, Atlanta, Georgia*, 1984.
- [Mata 05] M. T. Matache and J. Heidel. “Asynchronous random Boolean network model based on elementary cellular automata rule 126”. *Phys. Rev. E*, Vol. 71, p. 026232, 2005.
- [Matt 01] J. S. Mattick and M. J. Gagen. “The Evolution of Controlled Multitasked Gene Networks: The Role of Introns and Other Noncoding RNAs in the Development of Complex Organisms”. *Mol. Biol. Evol.*, Vol. 18, pp. 1611–1630, 2001.
- [McAd 95] H. H. McAdams and L. Shapiro. “Circuit Simulation of Genetic Networks”. *Science*, Vol. 269, No. 4, pp. 650–656, 1995.
- [McCu 43] W. S. McCulloch and W. Pitts. “A Logical Calculus of the Ideas Immanent in Nervous Activity”. *Bulletin of Mathematical Biophysics*, Vol. 5, pp. 115–133, 1943.
- [Meir 02] E. Meir, G. von Dassow, E. Munro, and G. M. Odell. “Robustness, Flexibility, and the Role of Lateral Inhibition in the Neurogenic Network”. *Current Biology*, Vol. 12, pp. 778–786, 2002.
- [Mend 98] L. Mendoza and E. Alvarez-Buylla. “Dynamics of the genetic network for *Arabidopsis thaliana* flower morphogenesis”. *J. Theor. Biol.*, Vol. 193, pp. 307–319, 1998.
- [Miha 06] T. Mihaljev and B. Drossel. “Scaling in a general class of critical random Boolean networks”. 2006. cond-mat/0606612.
- [Milo 04a] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, and U. Alon. “Response to Comment on Network Motifs: “Simple Building Blocks of Complex Networks” and “Superfamilies of Evolved and Designed Networks””. *Science*, Vol. 305, p. 1107, 2004.
- [Milo 04b] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat, M. Sheffer, and U. Alon. “Superfamilies of Evolved and Designed Networks”. *Science*, Vol. 303, pp. 1538–1542, 2004.
- [Mono 61] J. Monod and F. Jacob. “General conclusions: teleonomic mechanisms in cellular metabolism, growth, and differentiation”. *CSH Symp. Quant. Biol.*, Vol. 26, pp. 389–401, 1961.
- [Morg 47] O. Morgenstern and J. von Neumann. *The Theory of Games and Economic Behavior*. Princeton University Press, 1947.

- [Mura 96] E. Muraille, D. Thieffry, O. Leo, and M. Kauffman. “Toxicity and neuro-endocrine regulation of the immune response: a model analysis”. *J. Theor. Biol.*, Vol. 183, pp. 285–305, 1996.
- [Muti 04] M. A. Muñoz. “Multiplicative noise in non-equilibrium phase transitions: a tutorial”. In: E. Korutcheva and et al., Eds., *Advances in Condensed Matter and Statistical Physics*, p. 34, New York: Nova Science Publishers, 2004. cond-mat/0303650.
- [Nage 92] S. R. Nagel. “Instabilities in a sandpile”. *Rev. Mod. Phys.*, Vol. 64, p. 321, 1992.
- [Newm 06] M. E. J. Newman. “Modularity and community structure in networks”. *Proc. Nat. Acad. Sci.*, Vol. 103, No. 23, pp. 8577–8582, 2006.
- [Newm 99] M. E. J. Newman and R. G. Palmer. “Models of Extinction: A Review”. Working Papers of Santa Fe Institute, <http://ideas.repec.org/s/wop/safiw.html>, 1999.
- [Niko 06] S. Nikolajewa, M. Friedel, and T. Wilhelm. “Boolean Networks with biologically relevant rules show ordered behavior”. *BioSystems*, 2006. In press.
- [Odor 04] G. Odor. “Universality classes in nonequilibrium lattice systems”. *Rev. Mod. Phys.*, Vol. 76, p. 663, 2004.
- [Okte 03] H. Öktem, R. Pearson, and K. Egiazarian. “An adjustable aperiodic model class of genomic interactions using continuous time Boolean networks (Boolean delay equations)”. *Chaos*, Vol. 13, No. 4, p. 1167, 2003.
- [Ott 02] E. Ott. *Chaos in Dynamical Systems*. Cambridge University Press, New York, 2002.
- [Pasl 92] R. Paslack. “Ursprünge der Selbstorganisation”. In: S. J. Rusch, Gebhard/ Schmidt, Ed., *Konstruktivismus: Geschichte und Anwendung*, pp. 59–90, Suhrkamp, Frankfurt a.M., 1992.
- [Paul 06] U. Paul, V. Kaufman, and B. Drossel. “Properties of attractors of canalyzing random Boolean networks”. *Phys. Rev. E*, Vol. 73, p. 026118, 2006.
- [Pere 96] A. S. Perelson, A. U. Neumann, M. Markowitz, J. M. Leonard, and D. D. Ho. “HIV-1 dynamics in vivo: Virion clearance rate, infected cell lifespan, and viral generation time”. *Science*, Vol. 271, pp. 1582–1586, 1996.
- [Perk 06] T. J. Perkins, M. Hallett, and L. Glass. “Dynamical Modeling of Biological Regulatory Networks”. *Biosystems*, Vol. 84, No. 2, pp. 115–123, 2006.
- [Post 21] E. Post. “Introduction to a general theory of elementary propositions”. *Am. J. Math.*, Vol. 43, pp. 163–185, 1921.
- [Rao 02] C. V. Rao, D. M. Wolf, and A. P. Arkin. “Control, exploitation and tolerance of intracellular noise”. *Nature*, Vol. 420, pp. 231–237, 2002.

- [Rohl 02] T. Rohlf and S. Bornholdt. “Criticality in Random Threshold Networks: Annealed Approximation and Beyond”. *Physica A*, Vol. 310, pp. 245–259, 2002.
- [Ruvk 89] G. Ruvkun and J. Guisto. “The *Caenorhabditis elegans* heterochronic gene *lin-14* encodes a nuclear protein that forms a temporal developmental switch”. *Nature*, Vol. 338, pp. 313–319, 1989.
- [Samu 03] B. Samuelsson and C. Troein. “Superpolynomial growth in the number of attractors in Kauffman networks”. *Phys. Rev. Lett.*, Vol. 90, p. 098701, 2003.
- [Samu 05] B. Samuelsson and C. Troein. “Random maps and attractors in random Boolean networks”. *Phys. Rev. E*, Vol. 72, p. 046112, 2005.
- [Samu 06] B. Samuelsson and J. E. S. Socolar. “Exhaustive percolation on random networks”. *Phys. Rev. E*, Vol. 74, p. 036113, 2006.
- [Sanc 03] L. Sánchez and D. Thieffry. “Segmenting the fly embryo: a logical analysis of the pair-rule cross-regulatory module”. *J. Theor. Biol.*, Vol. 224, pp. 517–537, 2003.
- [Sanc 97] L. Sánchez, J. van Helden, and D. Thieffry. “Establishment of the dorso-ventral pattern during embryonic development of *Drosophila melanogaster*: a logical analysis”. *J. Theor. Biol.*, Vol. 189, pp. 377–389, 1997.
- [Schr 52] E. Schrödinger. *Science and Humanism. Physics in Our Time*. Cambridge: University Press, 1952.
- [Serr 01] R. Serra, M. Villani, and A. Salvemini. “Continuous genetic networks”. *Parallel Computing*, Vol. 27, pp. 663–683, 2001.
- [Shan 48] C. E. Shannon. “A Mathematical Theory of Communication”. *Bell System Technical Journal*, Vol. 27, pp. 379–423 & 623–656, 1948.
- [Shen 02] S. S. Shen-Orr, R. Milo, S. Mangan, and U. Alon. “Network motifs in the transcriptional regulation network of *Escherichia coli*”. *Nat. Genet.*, Vol. 31, pp. 64–68, 2002.
- [Shmu 02] I. Shmulevich, E. R. Dougherty, S. Kim, and W. Zhang. “Probabilistic Boolean networks: a rule - based uncertainty model for gene regulatory networks”. *Bioinformatics*, Vol. 18, No. 2, pp. 261–274, 2002.
- [Shmu 03] I. Shmulevich, H. Lähdesmäki, E. R. Dougherty, J. Astola, and W. Zhang. “The role of certain Post classes in Boolean network models of genetic networks”. *Proc. Nat. Acad. Sci.*, Vol. 100, No. 19, pp. 10734–10739, 2003.
- [Simo 01] I. Simon, J. Barnett, N. Hannett, C. T. Harbison, N. J. Rinaldi, T. L. Volkert, J. J. Wyrick, J. Zeitlinger, D. K. Gifford, T. S. Jaakkola, and R. A. Young. “Serial Regulation of Transcriptional Regulators in the Yeast Cell Cycle”. *Cell*, Vol. 106, pp. 697–708, 2001.

- [Simo 62] H. A. Simon. “The architecture of complexity”. In: *Proceedings of the American philosophical society*, pp. 467–482, 1962.
- [Smit 95] M. J. Smith and E. Szathmáry. *The Major Transitions in Evolution*. Oxford: Freeman Press, 1995.
- [Smut 26] J. Smuts. *Holism and Evolution*. New York: Macmillan Co., 1926.
- [Snou 89] E. H. Snoussi. “Qualitative dynamics of piece-linear differential equations: a discrete mapping approach”. *Dyn. Stability Syst.*, Vol. 4, pp. 189–207, 1989.
- [Snou 98] E. H. Snoussi. “Necessary Conditions for Multistationarity and Stable Periodicity”. *J. Biol. Syst.*, Vol. 6, pp. 3–9, 1998.
- [Soco 03] J. E. S. Socolar and S. A. Kauffman. “Scaling in ordered and critical random Boolean networks”. *Phys. Rev. Lett.*, Vol. 90, p. 068702, 2003.
- [Somo 96] R. Somogyi and C. A. Sniegoski. “Modeling the Complexity of Genetic Networks: Understanding Multigenic and Pleiotropic Regulation”. *Complexity*, Vol. 1, pp. 45–63, 1996.
- [Spir 01] P. Spirtes, C. Glymour, R. Scheines, S. Kauffman, V. Aimale, and F. Wimberly. “Constructing Bayesian network models of gene expression networks from microarray data”. In: *Proceedings of the Atlantic Symposium on Computational Biology, Genome Information Systems and Technology*, 2001.
- [Spor 02] O. Sporns and G. Tononi. “Classes of Network Connectivity and Dynamics”. *Complexity*, Vol. 7, No. 1, pp. 28–38, 2002.
- [Stau 87] D. Stauffer. “On forcing functions in Kauffman random Boolean networks”. *J. Stat. Phys.*, Vol. 46, No. 3-4, pp. 789–794, 1987.
- [Stau 94] D. Stauffer and A. Aharony. *Introduction to Percolation Theory*. CRC Press, 1994.
- [Steg 69] W. Stegmüller. *Main currents in contemporary German, British, and American philosophy*. Dordrecht, Reidel Publishing Co., 1969.
- [Steu 04] R. Steuer. “Effects of stochasticity in models of the cell cycle: from quantized cycle times to noise-induced oscillations”. *J. Theor. Biol.*, Vol. 228, pp. 293–301, 2004.
- [Swai 04] P. Swain. “Efficient attenuation of stochasticity in gene expression through posttranscriptional control”. *J. Mol. Biol.*, Vol. 344, p. 965, 2004.
- [Szal 98] Z. Szallasi and S. Liang. “Modeling the Normal and Neoplastic Cell Cycle with Realistic Boolean Genetic Networks: Their Application for Understanding Carcinogenesis and Assessing Therapeutic Strategies”. In: *Proc. Pacific Symp. on Biocomputing*, pp. 66–76, 1998.
- [Tana 06] R. J. Tanaka, H. Okano, and H. Kimura. “Mathematical Description of Gene Regulatory Units”. *Biophys J.*, Vol. 91, pp. 1235–1247, 2006.

- [Thie 95a] D. Thieffry, E. H. Snoussi, J. Richelle, and R. Thomas. “Positive loops and differentiation”. *J. Biol. Syst.*, Vol. 3, pp. 457–466, 1995.
- [Thie 95b] D. Thieffry and R. Thomas. “Dynamical behaviour of biological regulatory networks II. Immunity control in bacteriophage lambda”. *Bull. Math. Biol.*, Vol. 57, pp. 277–297, 1995.
- [Thie 98] D. Thieffry, A. M. Huerta, E. Perez-Rueda, and J. Collado-Vides. “From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in *Escherichia coli*”. *Bioessays*, Vol. 20, pp. 433–440, 1998.
- [Thie 99] D. Thieffry and D. Romero. “The modularity of biological regulatory networks”. *BioSystems*, Vol. 50, pp. 49–59, 1999.
- [Thom 01a] R. Thomas and M. Kaufman. “Multistationarity, the basis of cell differentiation and memory. I. Structural conditions of multistationarity and other nontrivial behavior”. *Chaos*, Vol. 11, No. 1, pp. 170–179, 2001.
- [Thom 01b] R. Thomas and M. Kaufman. “Multistationarity, the basis of cell differentiation and memory. II. Logical analysis of regulatory networks in terms of feedback circuits”. *Chaos*, Vol. 11, No. 1, pp. 180–195, 2001.
- [Thom 17] D. Thompson. *On Growth and Form*. Cambridge University Press, 1917. 1992 Dover Publications edition.
- [Thom 73] R. Thomas. “Boolean formalization of genetic control circuits”. *J. Theor. Biol.*, Vol. 42, pp. 563–585, 1973.
- [Thom 78] R. Thomas. “Logical analysis of systems comprising feedback loops”. *J. Theor. Biol.*, Vol. 73, pp. 631–656, 1978.
- [Thom 90] R. Thomas and R. D’Ari. *Biological Feedback*. CRC, Boca Raton, FL, 1990.
- [Thom 91] R. Thomas. “Regulatory networks seen as asynchronous automata: a logical description”. *J. Theor. Biol.*, Vol. 153, pp. 1–23, 1991.
- [Thom 95] R. Thomas, D. Thieffry, and M. Kaufman. “Dynamical behaviour of biological regulatory networks I: Biological role of feedback loops and practical use of the concept of the loop-characteristic state”. *Bull. Math. Biol.*, Vol. 57, pp. 247–276, 1995.
- [Tong 04] A. H. Y. Tong, G. Lesage, G. D. Bader, H. Ding, H. Xu, X. Xin, J. Young, G. F. Berriz, R. L. Brost, M. Chang, and et al. “Global Mapping of the Yeast Genetic Interaction Network”. *Science*, Vol. 303, pp. 808–813, 2004.
- [Turi 36] A. Turing. “On Computable Numbers, With an Application to the Entscheidungsproblem”. In: *Proceedings of the London Mathematical Society*, 1936.
- [Turi 52] A. Turing. “The chemical basis of morphogenesis”. *Phil. Trans. Roy. Soc. B*, Vol. 237, pp. 37–72, 1952.

- [Volk 98] L. G. Volkert and M. Conrad. “The role of weak interactions in biological systems: the dual dynamics model”. *J. Theor. Biol.*, Vol. 193, pp. 287–306, 1998.
- [Von 03] H. Von Foerster. *Understanding understanding : essays on cybernetics and cognition*. Springer: New York, Berlin, Heidelberg, Hong Kong, London, Milan, Paris, Tokyo, 2003.
- [Vose 99] M. D. Vose. *The Simple Genetic Algorithm: Foundations and Theory*. MIT Press, Cambridge, MA, 1999.
- [Watt 98] D. J. Watts and S. H. Strogatz. “Collective dynamics of small-world networks”. *Nature*, Vol. 393, p. 440, 1998.
- [Weav 99] C. Weaver, C. Workman, and G. Stormo. “Modeling regulatory networks with weight matrices”. Vol. 4, pp. 112–123, 1999.
- [West 06] R. Westra, G. Hollanders, R. Peeters, and K. Tuyls. “Piecewise Linear Modeling of Gene-Protein Interaction Networks”. In: *Proceedings of the Knowledge Discovery and Emergent Complexity workshop, Ghent, Belgium*, 2006.
- [Whit 86] J. G. White, E. Southgate, J. N. Thompson, and S. Brenner. “The structure of the nervous system of the nematode *C. elegans*”. *Philos. Trans. R. Soc. London Ser. B*, Vol. 314, pp. 1–340, 1986.
- [Wien 48] N. Wiener. *Cybernetics*. MIT Press, 1965 (1948).
- [Wilh 01] T. Wilhelm, J. Hollunder, J.-K. Kim, and A. Beyer. “Information Theoretic Measures for the Maturity of Ecosystems”. In: M. Matthies, H. Malchow, and J. Kriz, Eds., *Integrative Systems Approaches to Natural and Social Dynamics*, pp. 263–273, Springer, Berlin, 2001.
- [Wolf 02] S. Wolfram. *A New Kind of Science*. Wolfram Media Inc., New York, 2002.
- [Wolf 86] S. Wolfram. *Cellular Automata: Theory and Applications*. World Scientific Press, Singapore, 1986.
- [Yuh 01] C. H. Yuh, H. Bolouri, J. M. Bower, and E. H. Davidson. “A logical model of cis-regulatory control in a eukaryotic system”. In: J. M. Bower and H. Bolouri, Eds., *Computational Modeling of Genetic and Biochemical Networks*, pp. 73–100, MIT Press, Cambridge, MA, 2001.
- [Yuh 98] C. H. Yuh, H. Bolouri, and E. H. Davidson. “Genomic cis-regulatory logic: experimental and computational analysis of a sea urchin gene”. *Science*, Vol. 279, pp. 1896–1902, 1998.

Danksagung

Abschließend möchte ich nun allen danken, die mich auf dem Weg zur Fertigstellung meiner Dissertation begleitet, unterstützt und motiviert haben.

Begleitet, unterstützt und motiviert hat mich auf jeden Fall Frau Professor Dr. Barbara Drossel. Ohne ihre professionelle und liebevolle Art, auf die sie sich für das Gelingen dieser Arbeit eingesetzt hat, wäre diese Danksagung nicht zustande gekommen. Ich genoss die lockere, zur Wissenschaft anregende Atmosphäre in ihrer Arbeitsgruppe. Besonders möchte ich mich bei ihr für die vielen Möglichkeiten bedanken, mich außerhalb der TU Darmstadt persönlich und beruflich weiterzuentwickeln. Ich erinnere mich sehr gerne an die wissenschaftlichen Konferenzen und Tagungen, die Sommerschule, die Besuche von auswärtigen Wissenschaftlern, die Übungen, sowie an die Ausflüge und Musikabende.

Herrn Professor Dr. Porto danke ich für die Übernahme des Korreferats.

Herrn Professor Dr. Norbert Grewe und Herrn Priv. Doz. Dr. Frithjof Anders danke ich für alles, was sie mich gelehrt haben.

Unterstützt und motiviert haben mich zu verschiedenen Zeiten, durch Gespräche und Zusammenarbeit, ohne besondere Achtung auf die Reihenfolge, Herr Dr. Horst Turschner, Herr Dr. Timm Plefka, Herr Dr. Arijit Bhattacharyay, die Diplomphysiker und Diplomphysikerinnen Felix Wissel, Sebastian Schmitt, Frank Böhm, Torben Jabben, Satoshi Uchida, Tamara Mihaljev, Agnes Szejka, Ute Paul, Sebastian Weber, Florian Greil, Florian Teichert, sowie die Herrn Christian Guill und Christoph Fretter.

Meinen Freunden danke ich dafür, dass sie da waren. Meiner Frau danke ich für ihre Liebe und ihr Verständnis. Meinen Eltern danke ich für ihre Unterstützung.

Lebenslauf

Viktor Kaufman

- 04. Feb. 1976 Geboren in Lviv, Ukraine
- 1983 – 1991 Besuch der allgemeinbildenden Mittelschule Nummer 6 in Lviv, Ukraine
- 1991 – 1993 Besuch des Physik - Mathematik - Gymnasiums in Lviv mit Abschluss Attestat
- 1993 – 1997 Studium an der staatlichen I. Franko Universität in Lviv
- 1998 – 2001 Studium an der TU Darmstadt mit Abschluss Diplom
- seit Okt. 2001 Wissenschaftlicher Mitarbeiter am Institut für Festkörperphysik der TU Darmstadt

Eidesstattliche Erklärung:

Hiermit erkläre ich, dass ich die vorliegende Dissertation selbstständig verfasst, keine anderen als die angegebenen Hilfsmittel verwendet und bisher noch keinen Promotionsversuch unternommen habe.

Darmstadt, im Oktober 2006